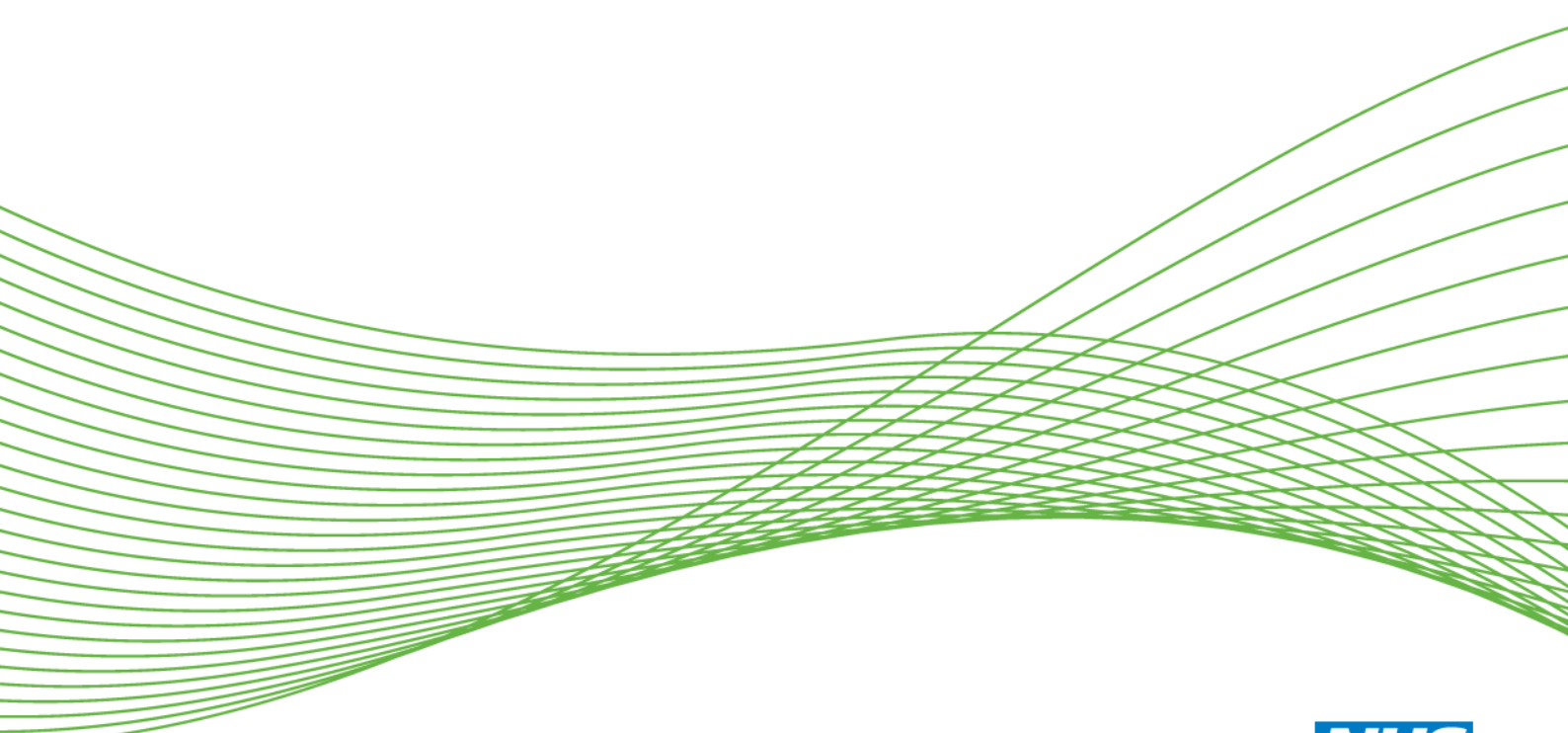


The ProFHER (PROximal Fracture of the Humerus: Evaluation by Randomisation) trial – a pragmatic multicentre randomised controlled trial evaluating the clinical effectiveness and cost-effectiveness of surgical compared with non-surgical treatment for proximal fracture of the humerus in adults

*Helen Handoll, Stephen Brealey, Amar Rangan, Ada Keding,
Belen Corbacho, Laura Jefferson, Ling-Hsiang Chuang,
Lorna Goodchild, Catherine Hewitt and David Torgerson*



**National Institute for
Health Research**

The ProFHER (PROximal Fracture of the Humerus: Evaluation by Randomisation) trial – a pragmatic multicentre randomised controlled trial evaluating the clinical effectiveness and cost-effectiveness of surgical compared with non-surgical treatment for proximal fracture of the humerus in adults

Helen Handoll,^{1*} Stephen Brealey,² Amar Rangan,³
Ada Keding,² Belen Corbacho,² Laura Jefferson,²
Ling-Hsiang Chuang,⁴ Lorna Goodchild,⁵
Catherine Hewitt² and David Torgerson²

¹Health and Social Care Institute, School of Health and Social Care,
Teesside University, Middlesbrough, UK

²York Trials Unit, Department of Health Sciences, University of York, York, UK

³Department of Trauma and Orthopaedics, James Cook University Hospital,
Middlesbrough, UK

⁴Pharmerit Europe, Rotterdam, the Netherlands

⁵Department of Rehabilitation, James Cook University Hospital,
Middlesbrough, UK

*Corresponding author

Declared competing interests of authors: Amar Rangan has obtained grants and personal fees from De Puy Ltd and grants from JRI Ltd; both are outside the submitted work. In addition, he has a UK and European patent application pending. None of these influenced the trial or the report.

This report should be referenced as follows:

Handoll H, Brealey S, Rangan A, Keding A, Corbacho B, Jefferson L, *et al.* The ProFHER (PROximal Fracture of the Humerus: Evaluation by Randomisation) trial – a pragmatic multicentre randomised controlled trial evaluating the clinical effectiveness and cost-effectiveness of surgical compared with non-surgical treatment for proximal fracture of the humerus in adults. *Health Technol Assess* 2015;**19**(24).

Health Technology Assessment is indexed and abstracted in *Index Medicus/MEDLINE*, *Excerpta Medica/EMBASE*, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®/Clinical Medicine*.

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 5.116

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index and is assessed for inclusion in the Database of Abstracts of Reviews of Effects.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: <http://www.nets.nihr.ac.uk/programmes/hta>

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 06/404/53. The contractual start date was in March 2008. The draft report began editorial review in December 2013 and was accepted for publication in June 2014. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2015. This work was produced by Handoll *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Editor-in-Chief of *Health Technology Assessment* and NIHR Journals Library

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the HTA Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andree Le May Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Professor Aileen Clarke Professor of Public Health and Health Services Research, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Peter Davidson Director of NETSCC, HTA, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Professor Elaine McColl Director, Newcastle Clinical Trials Unit, Institute of Health and Society, Newcastle University, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Health Sciences Research, Faculty of Education, University of Winchester, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Institute of Child Health, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: nihredit@southampton.ac.uk

Abstract

The ProFHER (PROximal Fracture of the Humerus: Evaluation by Randomisation) trial – a pragmatic multicentre randomised controlled trial evaluating the clinical effectiveness and cost-effectiveness of surgical compared with non-surgical treatment for proximal fracture of the humerus in adults

Helen Handoll,^{1*} Stephen Brealey,² Amar Rangan,³ Ada Keding,² Belen Corbacho,² Laura Jefferson,² Ling-Hsiang Chuang,⁴ Lorna Goodchild,⁵ Catherine Hewitt² and David Torgerson²

¹Health and Social Care Institute, School of Health and Social Care, Teesside University, Middlesbrough, UK

²York Trials Unit, Department of Health Sciences, University of York, York, UK

³Department of Trauma and Orthopaedics, James Cook University Hospital, Middlesbrough, UK

⁴Pharmerit Europe, Rotterdam, the Netherlands

⁵Department of Rehabilitation, James Cook University Hospital, Middlesbrough, UK

*Corresponding author H.Handoll@tees.ac.uk

Background: Proximal humeral fractures account for 5–6% of all fractures in adults. There is considerable variation in whether or not surgery is used in the management of displaced fractures involving the surgical neck.

Objective: To evaluate the clinical effectiveness and cost-effectiveness of surgical compared with non-surgical treatment of the majority of displaced fractures of the proximal humerus involving the surgical neck in adults.

Design: A pragmatic parallel-group multicentre randomised controlled trial with an economic evaluation. Follow-up was for 2 years.

Setting: Recruitment was undertaken in the orthopaedic departments of 33 acute NHS hospitals in the UK. Patient care pathways included outpatient and community-based rehabilitation.

Participants: Adults (aged ≥ 16 years) presenting within 3 weeks of their injury with a displaced fracture of the proximal humerus involving the surgical neck.

Interventions: The choice of surgical intervention was left to the treating surgeons, who used techniques with which they were experienced. Non-surgical treatment was initial sling immobilisation followed by active rehabilitation. Provision of rehabilitation was comparable in both groups.

Main outcome measures: The primary outcome was the Oxford Shoulder Score (OSS) assessed at 6, 12 and 24 months. Secondary outcomes were the 12-item Short Form health survey, surgical and other shoulder fracture-related complications, secondary surgery to the shoulder or increased/new shoulder-related therapy, medical complications during inpatient stay and mortality. European Quality of Life-5 Dimensions data and treatment costs were also collected.

Results: The mean age of the 250 trial participants was 66 years and 192 (77%) were female. Independent assessment using the Neer classification identified 18 one-part fractures, 128 two-part fractures and 104 three- or four-part fractures. OSS data were available for 215 participants at 2 years. We found no statistically or clinically significant differences in OSS scores between the two treatment groups (scale 0–48, with a higher score indicating a better outcome) over the 2-year period [difference of 0.75 points in favour of the surgery group, 95% confidence interval (CI) –1.33 to 2.84; $p = 0.479$; data from 114 surgery and 117 non-surgery participants] or at individual time points. We found no statistically significant differences between surgical and non-surgical group participants in SF-12 physical or mental component summary scores; surgical or shoulder fracture-related complications (30 vs. 23 respectively); those undergoing further shoulder-related therapy, either surgery (11 vs. 11 respectively) or other therapy (seven vs. four respectively); or mortality (nine vs. five respectively). The base-case economic analysis showed that, at 2 years, the cost of surgical intervention was, on average, £1780.73 more per patient (95% CI £1152.71 to £2408.75) than the cost of non-surgical intervention. It was also slightly less beneficial in terms of utilities, although this difference was not statistically significant. The net monetary benefit associated with surgery is negative. There was only a 5% probability of surgery achieving the criterion of costing < £20,000 to gain a quality-adjusted life-year, which was confirmed by extensive sensitivity analyses.

Conclusions: Current surgical practice does not result in a better outcome for most patients with displaced fractures of the proximal humerus involving the surgical neck and is not cost-effective in the UK setting. Two areas for future work are the setting up of a national database of these fractures, including the collection of patient-reported outcomes, and research on the best ways of informing patients with these and other upper limb fractures about initial self-care.

Trial registration: Current Controlled Trials ISRCTN50850043.

Funding: This project was funded by the NIHR Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 19, No. 24. See the NIHR Journals Library website for further project information.

Contents

List of tables	xv
List of figures	xxi
List of boxes	xxiii
List of abbreviations	xxv
Plain English summary	xxvii
Scientific summary	xxix
Chapter 1 Introduction	1
Proximal humeral fractures	1
<i>Fracture morphology, classification and epidemiology</i>	2
Treatment: care pathways	3
Evidence underpinning practice	4
Rationale for a trial comparing surgical with non-surgical treatment	4
Aim of the ProFHER trial	4
Chapter 2 Methods	7
Summary of trial design	7
Randomisation: sequence generation and allocation concealment	7
Participant eligibility	7
<i>Inclusion criteria</i>	8
<i>Exclusion criteria</i>	8
Sample population	8
Study setting	8
Interventions	8
<i>Surgery</i>	9
<i>Non-surgical intervention (control)</i>	9
<i>Rehabilitation</i>	9
Outcome measures	9
<i>Primary outcome measure: the Oxford Shoulder Score</i>	9
<i>Secondary outcomes</i>	10
<i>Data collection schedule</i>	10
<i>Data for the economic evaluation</i>	11
Classification of study fractures	11
Sample size	11
Estimating the numbers of patients and centres needed to recruit	12
<i>Proposed recruitment rate</i>	12
Governance	13
<i>Ethical arrangements</i>	13
<i>Trial governance</i>	13
Data handling	14
Data checks, cleaning and processing	14
Clinical effectiveness analysis	15
<i>Primary outcome analyses</i>	15

<i>Subgroup and sensitivity analyses</i>	15
<i>Secondary outcomes</i>	16
<i>Other outcomes</i>	16
<i>Lack of surgeon equipoise</i>	16
<i>The Neer classification of baseline radiographs</i>	16
Economic analysis	17
Chapter 3 Further methods: standardisation of trial procedures and care programmes, patient recruitment and data collection and processing	19
Standardisation of trial procedures and care programmes	19
<i>Site visits and information</i>	19
<i>Radiographs</i>	20
<i>Assessment of fractures</i>	20
<i>Provision of written patient information on personal care during initial sling immobilisation</i>	20
<i>Physiotherapy protocol</i>	20
<i>Home exercises</i>	21
<i>Delivery of physiotherapy</i>	21
Patient recruitment	22
<i>Identification of potential participants</i>	22
<i>Obtaining informed consent and randomisation</i>	22
Data collection	23
<i>Data handling and record keeping</i>	23
<i>Baseline data</i>	23
<i>Baseline radiographs</i>	23
<i>Patient questionnaires</i>	23
<i>Hospital forms</i>	25
Questionnaire on use of the ProFHER trial sling immobilisation leaflet and immobilisation	27
<i>Analysis of the questionnaire</i>	27
Data checks and 'cleaning': specific actions	27
Processing of data on complications and new therapy including surgery	28
<i>Surgical and other shoulder fracture-related complications, increased/new shoulder-related therapy and secondary surgery to the shoulder</i>	28
<i>Medical complications prior to discharge</i>	30
<i>Treatment for serious newly diagnosed medical complications</i>	30
<i>Further fractures</i>	30
<i>Adjusting the database for complications and treatment</i>	30
Adverse events classification	30
Baseline radiographs: processing and quality assessment	31
<i>Post-recruitment survey of hospital radiographers</i>	32
Preparations for the independent assessment and classification of radiographs based on the Neer classification	32
Chapter 4 Results: baseline characteristics of patients, delivery of interventions and return of questionnaires and other data	35
Patient recruitment	35
<i>Location in hospital of patients at recruitment</i>	35
Population sample	35
<i>Reasons for exclusion</i>	35
<i>Patient consent</i>	38
Participant flow	38
Baseline characteristics: all screened patients	40
Baseline characteristics of randomised patients	40
Interventions: surgery	44

<i>Operating surgeon</i>	44
<i>Summary of surgery data</i>	45
<i>Type of surgery/implant</i>	45
Interventions: non-surgical care	47
Findings from the questionnaire survey on sling immobilisation	48
<i>Use of the ProFHER trial sling immobilisation care leaflet</i>	48
<i>Type and duration of sling immobilisation typically provided to patients who would have been eligible for the ProFHER trial</i>	48
<i>Duration of sling immobilisation for non-surgical patients</i>	50
<i>Duration of sling immobilisation for post-surgical patients</i>	50
Physiotherapy	50
<i>Reasons for not starting physiotherapy</i>	52
<i>'Other' physiotherapy interventions</i>	52
<i>Home exercises</i>	53
<i>Referrals to other specialities</i>	53
<i>Reasons for any substantial differences from the ProFHER trial physiotherapy protocol</i>	53
<i>Provision of information leaflets illustrating home exercises</i>	53
<i>Completion of physiotherapy</i>	53
Return of data collection forms	54
<i>Patient questionnaires</i>	54
<i>Return of hospital data collection forms</i>	55
Other returns	56
<i>Baseline radiographs</i>	56
<i>Adverse events</i>	56
<i>Patient withdrawals</i>	56
Chapter 5 Primary outcome (Oxford Shoulder Score) analysis	57
Missing data	57
Descriptive Oxford Shoulder Score statistics	60
Primary analysis	61
Subgroup analyses	63
<i>Age</i>	63
<i>Fracture type</i>	63
<i>Independent Neer classification (sensitivity analysis)</i>	67
Sensitivity analysis: smoking	70
Patients' baseline treatment preferences	73
Exploration of centre variability	76
Summary of the Oxford Shoulder Score multilevel modelling results	76
Chapter 6 Results: analyses of secondary and other outcomes and supplementary analyses	81
12-item Short Form health survey	81
<i>Physical component summary score</i>	81
<i>Mental component summary score</i>	81
Surgical and shoulder fracture-related complications	86
Treatments for surgical and shoulder fracture-related complications	89
Medical complications during the inpatient stay	91
Other admissions and treatments: newly diagnosed medical complications and fractures	94
Mortality	97
Adverse events	97
<i>Serious adverse events</i>	97
<i>Non-serious adverse events</i>	98
<i>Expected adverse events and relatedness of adverse events to taking part in the trial</i>	99

'Untoward events' documented in the physiotherapy treatment logs	100
Discharge locations after inpatient stay	101
Patient treatment preferences at the end of the 2-year follow-up	101
Lack of surgeon equipoise	102
Surgeon preferences and agreed treatments for non-consenting patients	104
<i>Surgeon preferences</i>	104
<i>Agreed treatment for non-consenting patients</i>	104
Availability of radiographic views in the study population	105
Use of other imaging investigations	106
Independent assessment and classification of baseline radiographs according to the Neer classification	106
<i>Cull of radiographs in preparation for assessment</i>	106
<i>Assessment of available radiographic views</i>	106
<i>Assessment of radiograph quality</i>	107
<i>Fracture types and characteristics: agreement between the two raters</i>	108
<i>Agreed Neer classification of baseline fractures</i>	109
<i>Agreement between the baseline assessment of tuberosity involvement and the Neer classification</i>	111
Chapter 7 Economic evaluation	113
Objective	113
Overview	113
Methods	113
<i>Cost data sources</i>	113
<i>Health-related quality of life and quality-adjusted life-years</i>	114
<i>Health-care resource use and costs</i>	115
<i>Multiple imputation</i>	118
<i>Incremental analysis</i>	119
<i>Uncertainty</i>	120
Results	120
<i>Patient population</i>	120
<i>Health-related quality of life and quality-adjusted life-years</i>	120
<i>Health-care resource use and costs</i>	124
<i>Cost-effectiveness analysis and uncertainty</i>	127
<i>Sensitivity analysis</i>	129
Conclusion	130
Chapter 8 Discussion and conclusion	131
Summary of the main findings	131
<i>Primary outcome</i>	131
<i>Secondary outcomes</i>	132
<i>Cost-effectiveness</i>	133
Countering potential threats to internal validity	133
Applicability of the results	134
<i>Characteristics of the trial population</i>	134
<i>Applicability to the UK NHS</i>	136
Summary review of the currently available evidence	138
<i>Results</i>	139
<i>Pending and ongoing trials</i>	143
Application of the trial results to practice	143
Conclusion	144
<i>Recommendations for research</i>	144

Acknowledgements	147
References	149
Appendix 1 Hospital Episode Statistics for England: inpatient statistics (2003/4 to 2011/12) – totals and proximal humeral fracture	157
Appendix 2 Neer classification for proximal humeral fractures	159
Appendix 3 The ProFHER trial protocol version 7.0 (19 May 2009)	161
Appendix 4 Changes to the protocol	177
Appendix 5 Study eligibility form	181
Appendix 6 Sample population estimates	183
Appendix 7 Table of amendments sent to the Multicentre Research Ethics Committee	185
Appendix 8 The Trial Management Group and membership of the Trial Steering Committee and Data Monitoring and Ethics Committee	187
Appendix 9 Patient information sheet	189
Appendix 10 Payments for ProFHER trial-related activities	193
Appendix 11 Recommended radiographic views for the ProFHER trial	195
Appendix 12 The ProFHER trial sling immobilisation leaflet	199
Appendix 13 Questionnaire for physiotherapists on physiotherapy for proximal humeral fractures	203
Appendix 14 Physiotherapy protocol for the ProFHER trial	205
Appendix 15 Consent status form	209
Appendix 16 Patient consent form	211
Appendix 17 Baseline information form	213
Appendix 18 Three-month patient questionnaire	219
Appendix 19 Two-year patient questionnaire	225
Appendix 20 Treatment confirmation form	237
Appendix 21 Surgical form	239
Appendix 22 Inpatient episode form	241
Appendix 23 Physiotherapy treatment log	243

Appendix 24 Physiotherapy treatment log: completion of treatment ('end of treatment') form	245
Appendix 25 One-year follow-up form	247
Appendix 26 Two-year follow-up form	249
Appendix 27 Adverse event (reporting) form	251
Appendix 28 Review of adverse event form	253
Appendix 29 Trial exit form	255
Appendix 30 Questionnaire about sling care information	257
Appendix 31 Survey of radiography departments	259
Appendix 32 Post-recruitment survey of hospital radiographers	263
Appendix 33 The Neer classification of radiographs pro forma	267
Appendix 34 List of hospitals contributing to the ProFHER trial and primary care trusts for which research and development approval was obtained for follow-up	269
Appendix 35 Health-care collaborators at participating acute trust centres	271
Appendix 36 Report of the processing of complications and associated treatment including 'discards'	277
Appendix 37 Additional sensitivity analysis exploring the effect of clustering at centre level	279

List of tables

TABLE 1 Data collection schedule	11
TABLE 2 Sources of baseline characteristics data	25
TABLE 3 Descriptions of complications	29
TABLE 4 Coding scheme for 'outcome' of a complication or treatment for a complication	29
TABLE 5 Reasons listed for exclusion	37
TABLE 6 Baseline characteristics (different populations)	40
TABLE 7 Baseline characteristics (demographics) at randomisation and 24 months' follow-up	41
TABLE 8 Baseline characteristics (fracture data) at randomisation and 24 months' follow-up	42
TABLE 9 Baseline characteristics: types of fractures according to the Neer classification	44
TABLE 10 Staff grade of operating surgeons ($n = 109$ participants)	45
TABLE 11 Treatment details: surgical methods used in the surgery group	46
TABLE 12 Types of implant ($n = 109$)	47
TABLE 13 Prescribed non-surgical care (sling immobilisation) received by non-surgery patients	47
TABLE 14 Non-surgical treatments routinely/typically used by recruiting centres ($n = 32$)	48
TABLE 15 Post-surgical immobilisation routinely/typically used by recruiting centres ($n = 32$)	49
TABLE 16 Recruiting centres listing the same methods for non-surgical and post-surgical immobilisation	50
TABLE 17 Physiotherapy treatment	51
TABLE 18 Completion of physiotherapy	54
TABLE 19 Return of patient questionnaires	54
TABLE 20 Days from due follow-up date to completion (date recorded on the patient questionnaire) and return (logged by the YTU) of the questionnaires	55

TABLE 21 Return of hospital forms	55
TABLE 22 Missing primary outcome data over time	57
TABLE 23 Predictors of non-response: demographics	58
TABLE 24 Predictors of non-response: fracture data	59
TABLE 25 Predictors of non-response: OSS total score	60
TABLE 26 Unadjusted descriptive OSS statistics over time by treatment group	60
TABLE 27 Difference in mean OSS scores over time by treatment group: primary analysis using a multilevel model	63
TABLE 28 Descriptive OSS statistics over time by treatment group and age group	64
TABLE 29 Difference in mean OSS scores over time by treatment group: subgroup analysis 1 – age (< 65 years/≥ 65 years)	65
TABLE 30 Descriptive OSS statistics over time by treatment group and tuberosity involvement	66
TABLE 31 Difference in mean OSS scores over time by treatment group: subgroup analysis 2 – tuberosity involvement (yes/no)	68
TABLE 32 Descriptive OSS statistics over time by treatment group and Neer classification	68
TABLE 33 Difference in mean OSS scores over time by treatment group: subgroup analysis 2 – Neer grouping (one- and two-part fractures/three- and four-part fractures)	70
TABLE 34 Descriptive OSS statistics over time by treatment group and smoking status	71
TABLE 35 Difference in mean OSS scores over time by treatment group: sensitivity analysis – smoking	72
TABLE 36 Treatment preference of eligible patients at baseline	73
TABLE 37 Descriptive OSS statistics over time by treatment group and patient preference	74
TABLE 38 Difference in mean OSS scores over time by treatment group: including patient preferences at baseline	76
TABLE 39 Difference in mean OSS scores over time by treatment group	79
TABLE 40 Descriptive SF-12 PCS score statistics over time by treatment group	82
TABLE 41 Difference in mean SF-12 PCS scores over time by treatment group	84

TABLE 42 Descriptive SF-12 MCS score statistics over time by treatment group	84
TABLE 43 Difference in mean SF-12 MCS scores over time by treatment group	86
TABLE 44 Numbers of surgical and shoulder fracture-related complications by treatment group	87
TABLE 45 Baseline characteristics by presence and absence of surgical and shoulder fracture-related complications	88
TABLE 46 Number of secondary surgeries or increased/new therapies by treatment group	90
TABLE 47 Shoulder fracture-related complications: treatments and outcomes by treatment group	90
TABLE 48 Details of the treatments and outcomes associated with the 36 complications in the surgery group	91
TABLE 49 Details of the treatments and outcomes associated with the 23 complications in the non-surgery group	93
TABLE 50 Number of medical complications during the inpatient stay by treatment group	94
TABLE 51 Numbers of other admissions and treatments by treatment group	95
TABLE 52 Further fractures by treatment group	97
TABLE 53 Mortality by treatment group	97
TABLE 54 Numbers of patients experiencing one or more SAE by treatment group	98
TABLE 55 Numbers of SAEs by category and treatment group	98
TABLE 56 Numbers of patients experiencing one or more non-SAE by treatment group	98
TABLE 57 Numbers of non-SAEs by category and treatment group	99
TABLE 58 Expected adverse events and relatedness of adverse events to trial participation by treatment group	99
TABLE 59 Additional complications reported in the physiotherapy logs for surgical group patients	100
TABLE 60 Additional complications reported in the physiotherapy logs for non-surgical group patients	100
TABLE 61 Treatment preferences of trial participants at baseline and 24 months	101
TABLE 62 Treatment preferences of trial participants at 24 months by treatment group	101

TABLE 63 Baseline characteristics by lack of equipoise exclusion	103
TABLE 64 Baseline characteristics by surgeon advice for patients excluded because of a lack of equipoise	104
TABLE 65 Treatment advice by clinicians for patients who were not entered into the trial	104
TABLE 66 Preferences and agreed treatments for eligible patients who did not consent	105
TABLE 67 Radiographic views recorded on the study eligibility forms	105
TABLE 68 Radiographic views available for study participants (eligibility forms and Neer classification proforma)	107
TABLE 69 Agreement on radiograph quality criteria between raters	107
TABLE 70 Neer classification: initial assignments by the two raters acting independently	108
TABLE 71 Separate fracture characteristics on the Neer classifications proforma: individual judgements of raters	109
TABLE 72 Agreed Neer classification by treatment group	109
TABLE 73 'Unexpected' Neer categories: involvement/displacement of anatomical structures (Neer classification proforma)	110
TABLE 74 'Unexpected' Neer categories: involvement of the tuberosities (Neer classification proforma plus eligibility forms)	111
TABLE 75 Agreed Neer classification and tuberosity assessment at baseline: overview	111
TABLE 76 Agreed Neer classification and tuberosity assessment at baseline: detail	112
TABLE 77 Unit costs (and sources) used to estimate the total cost for each individual patient: primary surgery	116
TABLE 78 Unit costs (and sources) used to estimate total cost for each individual patient: primary care, secondary care and physiotherapy	117
TABLE 79 Health-related quality of life: numbers of questionnaires returned and completed EQ-5D scores and corresponding proportions of missing data by allocation and follow-up	121
TABLE 80 Health-related quality of life: EQ-5D scores at baseline and at follow-up points up to 2 years	121
TABLE 81 Health-related quality of life: total QALYs for all available cases by allocation over the 2 years of follow-up	124

TABLE 82 Average resource use per arm of the trial over the 2 years of follow-up	125
TABLE 83 Cost by category and treatment allocation	126
TABLE 84 Comparison of mean total costs per allocation during follow-up periods	126
TABLE 85 Summary of the incremental analysis (ITT), cost-effectiveness results and uncertainty for the base case	127
TABLE 86 Sensitivity analysis: summary of the incremental analysis (ITT), cost-effectiveness results and uncertainty when considering all resource use	129
TABLE 87 Sensitivity analysis: summary of the incremental analysis (ITT), cost-effectiveness results and uncertainty when considering patients' questionnaires	130
TABLE 88 Patient characteristics in participating and non-participating trusts	138
TABLE 89 Characteristics of the eight RCTs comparing surgery with conservative treatment for proximal humeral fractures	139
TABLE 90 Ongoing trials	143
TABLE 91 Views recommended by hospital radiographers for patients with a proximal humeral fracture ($n = 26$)	263
TABLE 92 Radiographer agreement with quality criteria ($n = 26$)	264
TABLE 93 Difference in mean OSS scores over time by treatment group: sensitivity analysis – adjusted for clustering by centre	279

List of figures

FIGURE 1 Proximal humerus anatomy and the four segments	2
FIGURE 2 Two-part fracture of (a) the surgical neck with (b) subsequent internal fixation with plate and screws	3
FIGURE 3 Flow chart of patient recruitment and hospital data collection	24
FIGURE 4 Patient recruitment and accrual in the ProFHER trial	36
FIGURE 5 Patient recruitment by centre listed in order of when the centre was set up to recruit	37
FIGURE 6 Non-consenting and consenting patients by centre	38
FIGURE 7 Participant flow through the trial	39
FIGURE 8 Mean OSS scores (with 95% confidence intervals) over time by treatment group	61
FIGURE 9 Distribution of OSS scores over time	62
FIGURE 10 Mean OSS scores (with 95% CIs) over time by treatment group and age group	65
FIGURE 11 Mean OSS scores (with 95% CIs) over time by treatment group and tuberosity involvement	67
FIGURE 12 Mean OSS scores (with 95% CIs) over time by treatment group and Neer grouping	69
FIGURE 13 Mean OSS scores (with 95% CIs) over time by treatment group and smoking status	72
FIGURE 14 Mean OSS scores (with 95% CIs) over time by treatment group and patient preference	75
FIGURE 15 Mean OSS scores (with 95% CIs) over time by recruitment site	77
FIGURE 16 Distribution of OSS scores (box plots) by recruitment site	78
FIGURE 17 Mean SF-12 PCS scores (with 95% CIs) over time by treatment group	82
FIGURE 18 Distribution of SF-12 PCS scores over time	83
FIGURE 19 Mean SF-12 MCS scores (with 95% CIs) over time by treatment group	84
FIGURE 20 Distribution of SF-12 MCS scores over time	85
FIGURE 21 Lack of surgeon equipoise and ineligible patients by centre	102

FIGURE 22 Pre- and post-cull radiographic views per patient	106
FIGURE 23 Distribution of EQ-5D scores at baseline and follow-up points up to 2 years	122
FIGURE 24 Mean EQ-5D scores at baseline and follow-up points up to 2 years	123
FIGURE 25 Cost-effectiveness plane, controlling for covariates: base-case analysis – non-parametric bootstrapping results	128
FIGURE 26 Confidence ellipses, controlling for covariates: base-case analysis	128
FIGURE 27 Cost-effectiveness acceptability curve, controlling for covariates: base-case analysis	129
FIGURE 28 Four representative fractures from the ProFHER trial	135
FIGURE 29 Average number of FCEs in participating and non-participating trusts	137
FIGURE 30 Average number of admissions in participating and non-participating trusts	137
FIGURE 31 Meta-analysis of patient-reported functional scores	140
FIGURE 32 Meta-analysis of subsequent shoulder surgery	141
FIGURE 33 Meta-analysis of mortality	142

List of boxes

BOX 1 Categories for coding adverse events	31
BOX 2 Unexpected surgical procedures	47

List of abbreviations

A&E	accident and emergency	MI	myocardial infarction
ADL	activities of daily living	MREC	Multicentre Research Ethics Committee
AUC	area under the curve	NICE	National Institute for Health and Care Excellence
AVN	avascular necrosis	NMB	net monetary benefit
BNF	<i>British National Formulary</i>	OLS	ordinary least squares
CEAC	cost-effectiveness acceptability curve	OSS	Oxford Shoulder Score
CI	confidence interval	PACS	Picture Archiving and Communications Systems
CONSORT	Consolidated Standards of Reporting Trials	PCS	physical component summary
CRF	case report form	ProFHER	PROximal Fracture of the Humerus: Evaluation by Randomisation
CT	computerised tomography	PSSRU	Personal Social Services Research Unit
CV	curriculum vitae	QALY	quality-adjusted life-year
DMEC	Data Monitoring and Ethics Committee	R&D	research and development
DVT	deep-vein thrombosis	RCT	randomised controlled trial
EQ-5D	European Quality of Life-5 Dimensions	SAE	serious adverse event
FCE	finished consultant episode	SD	standard deviation
GP	general practitioner	SF-12	12-item Short Form health survey
HES	Hospital Episode Statistics	SF-36	36-item Short Form health survey
HRG	Healthcare Resource Group	TENS	transcutaneous electrical nerve stimulation
HRQoL	health-related quality of life	TMG	Trial Management Group
ICER	incremental cost-effectiveness ratio	TSC	Trial Steering Committee
IQR	interquartile range	YTU	York Trials Unit
ITT	intention to treat	WTP	willingness to pay
JPEG	Joint Photographic Experts Group		
MAR	missing at random		
MCS	mental component summary		

Plain English summary

Fracture of the proximal humerus (the top part of the upper arm bone) is common, particularly in older adults. Generally, less serious fractures are successfully treated by supporting the injured arm in a sling while the fracture mends. More serious fractures, in which two or more bony fragments are displaced, are often treated surgically. This usually involves putting the fracture fragments back in place and fixing them with plates and screws. In more complex fractures, the humeral head is sometimes replaced by an artificial joint. However, we identified a need for evidence to clarify whether surgery or non-surgical treatment is best for the more common types of displaced fracture.

We conducted a randomised controlled trial in 35 NHS hospitals that compared surgery with non-surgical treatment for these fractures. Of the 250 patients who enrolled into the trial, 125 were randomised to 'surgery' and 125 to 'not surgery'. Questionnaires, which collected data on patients' shoulder function and general health at 6, 12 and 24 months, were received back from 215 participants at 24 months' follow-up. Analysis of these questionnaires and the forms from hospitals showed that there were no important differences in shoulder function, general health, complications or treatment for complications between patients in the surgery group and those in the non-surgery group. Analysis of the treatment costs incurred over the 2 years showed that surgery also costs more. In summary, the evidence from this trial shows that current methods of surgery do not result in a better outcome for these patients and are more costly.

Scientific summary

Background

Proximal humeral fractures occur mainly in older adults and account for 5–6% of all fractures in adults. It is estimated that the substantial costs associated with these fractures are about one-fifth of those of hip fracture in the first year. Approximately 40% of these fractures are displaced fractures involving the surgical neck. As reflected in the variation in practice, the role of surgery for this group of fractures is unclear but its use is increasing.

Objectives

The primary objective was to evaluate the clinical effectiveness and cost-effectiveness of surgery compared with non-surgical treatment of the majority of displaced fractures of the proximal humerus involving the surgical neck in adults.

Secondary objectives included:

- to describe the study population in terms of key characteristics and in particular to categorise the trial fractures according to the Neer classification
- to ensure that both groups of patients received comparable and good standards of care
- to explore the effect of age (< 65 years vs. ≥ 65 years), type of fracture (involving none vs. one or both tuberosities) and patients' treatment preference on the primary outcome [Oxford Shoulder Score (OSS)].

Design

The PROximal Fracture of the Humerus: Evaluation by Randomisation (ProFHER) trial was a pragmatic parallel-group multicentre randomised controlled trial (RCT) with an economic evaluation. Patients were randomised on an equal basis to receive surgical or non-surgical treatment via a remote randomisation service. Randomisation was stratified by tuberosity involvement (yes or no) with allocation using random block sizes. Follow-up was for 2 years. Data collection, which included copies of baseline radiographs for all randomised patients, was by means of hospital and physiotherapy forms, and patient questionnaires. There was no blinding of outcome assessment. However, the independent classification of the baseline radiographs and all coding of data, which involved at least two independent raters, was carried out blind to treatment allocation.

Throughout the conduct of the trial we emphasised the importance of good practice from clinical and research perspectives; standardised protocols and care pathways; and comparable and sufficient expertise of care providers in the two groups. As far as possible we adopted procedures that reflected and were compatible with usual practice and that did not delay treatment, add unnecessarily to the clinical workload or place burden on participants.

Setting

Trial recruitment was undertaken in the orthopaedic departments of 33 of the 35 participating acute NHS hospitals; 33 participating hospitals were based in England and one each was based in Scotland and Wales. Patient care pathways included outpatient and community-based rehabilitation. Patients were recruited from fracture clinics or orthopaedic wards from 17 September 2008 up to 13 April 2011.

Participants

Adults (aged ≥ 16 years) presenting to the participating hospital within 3 weeks of their injury with a radiographically confirmed displaced fracture of the proximal humerus involving the surgical neck were considered for inclusion. Fractures that were displaced enough for the treating surgeon to consider surgical intervention were considered for inclusion, even if the degree of displacement did not meet the exact displacement criteria of the Neer classification (1 cm and/or 45° angulation of displaced parts). (We refer to fractures not meeting these criteria as 'one-part' fractures.) Exclusion criteria were associated dislocation of the injured shoulder joint; open fracture; lack of mental capacity to understand the trial or instructions for rehabilitation; comorbidities precluding surgery/anaesthesia; clear indication for surgery such as severe soft-tissue compromise; multiple injuries (upper limb fractures); pathological fractures (other than osteoporotic) and terminal illness; and participant not resident in the hospital catchment area. Patients were provided with an ethically approved information sheet and written consent was obtained before randomisation.

Interventions

The choice of surgical intervention was left to the treating surgeons, who used techniques that they were fully experienced with to avoid learning curve problems. Non-surgical treatment was initial sling immobilisation followed by active rehabilitation. Measures taken to ensure comparable delivery of care and rehabilitation for both groups of patients included the development and provision of an information leaflet on personal care during initial sling immobilisation, a physiotherapy protocol to promote standard rehabilitation care and promotion of the need for patients to perform home exercises.

Main outcome measures

The primary outcome measure was the OSS (scale 0–48, with higher scores indicating a better outcome) assessed at 6, 12 and 24 months. The trial was powered to detect a clinically important difference in the OSS of 5 points (equivalent to an effect size of 0.4).

Secondary outcomes were the 12-item Short Form health survey (SF-12) [physical and mental component summary scores (PCS; MCS) scale 0–100, with higher scores indicating better health], surgical and other shoulder fracture-related complications, secondary surgery to the shoulder or increased/new shoulder-related therapy, medical complications during the inpatient stay and mortality. European Quality of Life-5 Dimensions (EQ-5D) data and treatment costs incurred in hospital and subsequently were also collected.

Statistical analysis

Both the clinical effectiveness and economic analyses were conducted to a prespecified and externally endorsed analysis plan. Intention-to-treat analysis was conducted throughout.

The primary clinical effectiveness analysis compared the results of the two groups over all three follow-up assessments. A random slope multilevel model was fitted to the data with time points nested in patients to allow for clustering of data within each patient. This model adjusted for the fixed effects of treatment group, time (6, 12 or 24 months), the interaction between treatment and time, tuberosity involvement at baseline (yes or no), age (< 65 years or ≥ 65 years), gender and health status at baseline (EQ-5D). Different covariance patterns for the repeated measurements were explored and the most appropriate pattern selected for the final model. Estimates of the difference in OSS between treatment groups were assessed overall and at individual time points. All main analyses were performed using two-sided significance tests at the 5% significance level.

Two subgroup analyses, whose expected direction of effect was specified a priori, were planned for the OSS only. These were based on age (< 65 years or ≥ 65 years), with younger patients expected to benefit more from surgery, and tuberosity involvement (yes or no) recorded by the surgeon who assessed patient eligibility at baseline, with patients expected to benefit from surgery over non-surgical treatment when one or both tuberosities were involved. A sensitivity analysis of tuberosity involvement in terms of Neer one- and two-part fractures compared with Neer three- and four-part fractures was undertaken. This was based on the classification of the fractures by two independent experts using baseline radiographs. The primary analysis model was extended for each of these subgroup analyses by including, in turn, an interaction of treatment group with age group, tuberosity involvement or Neer parts. The differences in 2-log likelihood model fit between these and the base model were assessed using the chi-squared test at $p < 0.10$.

Reflecting a chance baseline imbalance, smoking status was added to the primary model as a sensitivity analysis. To ascertain the impact of prior beliefs, an interaction of treatment group with patient treatment preference was added to the base model. The variability in the OSS across recruitment sites was analysed descriptively.

The SF-12 PCS and MCS scores were analysed by multilevel modelling using the same predictors as in the OSS primary model. Frequencies of shoulder surgery- and fracture-related complications and any treatments for these were compared between treatment groups using the chi-squared test. Mortality rates were reported and compared between treatment groups using the chi-squared test.

Economic analysis

The economic analysis was carried out from the UK NHS and Personal Social Services perspective. Data on health utilities, obtained from the EQ-5D data collected from patient questionnaires, were converted into quality-adjusted life-years (QALYs) for each patient using the area under the curve method. Costs were expressed in UK pounds sterling at a 2012 price base.

Differences in mean costs and QALYs at 2 years were used to derive an estimate of the cost-effectiveness of surgery and non-surgical treatment. Multiple imputation was used to derive the data set for the base-case analysis. An additional analysis was conducted for complete cases (in which patients with any missing data are excluded). We used sensitivity analyses to explore the impact of incorporating both types of resource use (shoulder and non-shoulder related) and using different data sources. Cost-effectiveness acceptability curves (CEACs) were used to express the probability of whether or not surgery is cost-effective at the willingness-to-pay threshold.

Clinical effectiveness results

Of 1250 patients screened, 563 were considered eligible, of whom 250 were recruited. The mean age of the trial participants was 66 years and 192 (77%) were female. Independent characterisation of radiographs confirmed that the patients included in the ProFHER trial had sustained injuries that are typically considered for surgical intervention in contemporaneous practice. The assessment, based on the Neer classification, identified 18 one-part fractures, 128 two-part fractures and 104 three- or four-part fractures.

Of the 125 patients allocated to surgery, 109 received surgery. Consultant surgeons were predominantly involved in the management of these fractures and usually performed the operations. The choice of implants, with the majority being locking plates and a minority receiving hemiarthroplasty, is compatible with expectations of current practice. Of the 125 patients allocated non-surgical treatment, 123 received conservative care. Good and comparable non-surgical care in the two groups was achieved as evidenced

by (1) the provision of the ProFHER trial sling information leaflet to nearly all trial patients; (2) the equivalent provision of physiotherapy in terms of timing, numbers of sessions and interventions applied; and (3) comprehensive recording by physiotherapists showing that the majority of patients were doing their home exercises.

The treatment groups were balanced for all baseline characteristics, including Neer categories, except for smoking status. These profiles also applied to the baseline characteristics of the 215 patients (106 surgery patients vs. 109 non-surgery patients) with OSS data at 2 years.

The primary analysis included data from 231 patients (114 surgery patients vs. 117 non-surgery patients) with valid OSS data for at least one follow-up time point and complete baseline covariates. We found that there were no statistically significant differences in OSS between the two treatment groups over the 2-year period [difference of 0.75 points in favour of the surgery group, 95% confidence interval (CI) –1.33 to 2.84; $p = 0.479$] or at individual time points. All 95% CIs excluded the prespecified difference in OSS of 5 points, thus indicating that any observed difference was not of 'clinical significance'.

There was no statistically significant effect of treatment group when including interactions with age [comparison with base model: $\chi^2(1) = 1.24$, $p = 0.265$] or fracture type [assessed by tuberosity involvement at baseline; $\chi^2(1) < 0.01$, $p = 0.954$] and Neer classification [$\chi^2(1) = 0.05$, $p = 0.823$] in the two planned subgroup analyses. Thus, our prior expectations of directions of effect (subgroup differences) were not supported by these results and strengthen the case for not differentiating treatment (use of surgery) on the basis of these characteristics. Similarly, the effect of treatment group remained not statistically significant when accounting for smoking status and patient treatment preference.

We found no statistically significant differences between treatment groups in the SF-12 PCS score. The PCS score was on average 1.8 score points higher in the surgical group than in the non-surgical group (95% CI –0.84 to 4.39, $p = 0.184$). The same lack of significance applied for the SF-12 MCS score. The MCS score was on average 1.3 points lower in the surgical group than in the non-surgical group (95% CI –3.80 to 1.23, $p = 0.317$).

All 10 early medical complications occurred in surgery group participants. Slightly more patients in the surgery group than in the non-surgery group experienced a surgical or shoulder fracture-related complication [30 (24%) vs. 23 (18%)]. Although the same number in each group ($n = 11$; 9%) had secondary surgery to the shoulder within the 2-year follow-up period, slightly more in the surgery group than in the non-surgery group had increased or new shoulder-related therapy [seven (5.6%) vs. four (3.2%)]. Neither of these differences were statistically significant (patients with complications $p = 0.279$; patients undergoing further surgery or therapy $p = 0.575$). The same finding applied to mortality, with slightly more deaths in the surgery group than in the non-surgery group [nine (7.2%) vs. five (4.0%), $p = 0.271$]. In total, 28 patients in each group experienced at least one serious adverse event (SAE). Nine patients experienced at least one non-SAE, of whom four were in the surgery group and five were in the non-surgery group.

Cost-effectiveness results

The base-case economic analysis showed that, at 2 years, the cost of surgical intervention was, on average, £1780.73 more per patient (95% CI £1152.71 to £2408.75) than the cost of non-surgical intervention. It was also slightly less beneficial in terms of utilities, although this difference was not statistically significant (difference of –0.0158, 95% CI –0.13 to 0.10 when adjusted for baseline utility; –0.0101, 95% CI –0.13 to 0.11 when adjusted for covariates). The net monetary benefit associated with surgery is negative, indicating that the resources to be displaced would be greater than the benefit to be gained if surgery were implemented in the NHS. Furthermore, the CEAC showed that surgery had only a 5% probability of achieving the criterion of costing < £20,000 to gain a QALY. Therefore, surgery appears to

be a dominated treatment option and not a cost-effective use of health-care resources. These findings were robust to the three sensitivity analyses that were undertaken. Although surgery did not result in a dominated alternative for the base-case analysis or the sensitivity analyses that included both shoulder and non-shoulder-related resource use, the incremental cost-effectiveness ratios were still above National Institute for Health and Care Excellence cost-effectiveness thresholds (£20,000–30,000 per QALY gained).

Conclusion

The ProFHER trial has provided robust clinically relevant evidence showing that current surgical practice does not result in a better outcome for most patients with a displaced fracture of the proximal humerus involving the surgical neck and that, in addition, it is not cost-effective in the UK setting.

It is important that non-surgical care should be of a good standard, including the availability of a leaflet about sling immobilisation, timely access to physiotherapy and promotion of home exercises. The potential need for remedial surgery for severe symptomatic complications in around 5–10% of these patients should be factored into forecasts for hospital budgets.

Recommendations for research

Given the above findings and the existence of five ongoing trials, initiating further RCTs on this question is not appropriate. The setting up of a national database of these fractures, with the systematic and prospective collection of data on epidemiology, management and outcome, including of patient-reported outcomes, should be considered. In addition, research is required to establish the best approach to providing patient information on the early treatment of these fractures.

Trial registration

This trial is registered as ISRCTN50850043.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1 Introduction

Proximal humeral fractures

Fractures of the proximal humerus are those occurring at the top end of the humerus or upper arm bone. An estimated 706,000 proximal humeral fractures occurred worldwide in 2000.¹ In the same year, an incidence of 63 proximal humeral fractures per 100,000 adults was reported for the population (534,715 adults) served by the Edinburgh Royal Infirmary,² and an incidence of 79 humeral fractures, of which 61 were likely to be proximal humeral fractures based on fracture distribution statistics,² per 100,000 adults aged > 20 years was reported in England and Wales.³ Based on 2001 UK census data, this amounted to an estimated 23,788 proximal humeral fractures in England and Wales (population 38,997,087 aged > 20 years).⁴ These fractures accounted for 5–6% of all adult fractures in studies conducted in Scotland² and Denmark.⁵ Their incidence increases rapidly with age, with the majority occurring in people aged > 65 years.^{2,5} Women are affected between two and three times as often as men.^{2,5,6} Most of these fractures result from low-energy trauma, predominantly a fall from standing height.^{5,7} The primary injury mechanism, the gender ratio and the age distribution point to underlying osteoporosis.^{7,8} Similar to other primarily osteoporotic fractures, the incidence and age-specific incidence of these fractures are increasing.⁸

The immediate and long-term consequences of proximal humeral fractures are substantial. Especially in the elderly, who will often have other comorbidities, these fractures can result in lengthy inpatient rehabilitation and, for some, a loss of independence such as a permanent move to a nursing home.⁹ Hospital Episode Statistics (HES) data for England¹⁰ over a 9-year period (2003–12) consistently show a mean inpatient stay of > 1 week and that just over half of these episodes relate to patients aged ≥ 75 years (see *Appendix 1*). The outcome following treatment of these fractures is frequently unsatisfactory. Poor shoulder function and pain are common long-term outcomes, with a substantial proportion of patients, even those with less severe injuries, continuing to report some or severe disability at 2 years.¹¹ These injuries are also associated with an increase in mortality.¹²

Although dwarfed by the costs of hip (proximal femur) fracture, for which the bulk of specific economic cost data on osteoporotic fractures are available, the costs of proximal humeral fractures are still considerable and rising. One study reporting data collected from 1989 to 1991 on residents of Olmsted County, MN, USA, reported a 0.21 cost ratio of treating a humeral fracture compared with a hip fracture in the first year.¹³ Again in the USA, a retrospective study of fracture-related direct medical costs in 2000 found that the cost per fracture per year for a humeral fracture was 21% that of a hip fracture (US\$5567 vs. US\$26,856 at 2003 Medicare fee schedule payment levels).¹⁴ Given the 3.2% increase in the surgical management of proximal humeral fractures (most hip fractures are managed surgically) between 1999–2000 and 2004–5, it is likely that the relative cost of these fractures has increased in the USA.¹⁵ A report providing direct health-care costs of injuries to the shoulder, arm and wrist in the Netherlands, adjusted to 2007 prices, reported that upper arm fractures (annual average of 9038 proximal humeral and humeral shaft fractures) were the most expensive injuries per case (€4440) and that their overall annual cost was approximately €40M out of an estimated €290M cost for upper extremity injuries.¹⁶ It was suggested that the increase in the cost of fracture care in 'elderly women' from a previous report of costs in the Netherlands was partly because of a higher number of patients receiving operations for these fractures.¹⁶

Fracture morphology, classification and epidemiology

Over 99% of proximal humeral fractures are closed fractures in which the overlying skin remains intact.¹⁷ The most commonly used classification of these fractures is that of Neer.¹⁸ Neer based his 16-category classification primarily on the relative positions of the four main segments of the proximal humerus: the humeral head, the greater tuberosity, the lesser tuberosity and the humeral shaft (see *Figure 1*). Although these may be delineated by fracture lines, a segment is only considered a 'part' if there is displacement of > 1 cm or 45° angulation. A 'minimally displaced' or 'undisplaced' fracture occurs when the displacement criteria are not met for any of the four segments. For ease of description and to avoid the misinterpretation that these are all 'undisplaced' or 'minimally displaced' fractures, in common with other researchers,¹⁹ we generally refer to such fractures as one-part fractures in this report. The other categories – two-part, three-part and four-part fractures – involve the relative displacement of two, three or all four segments respectively. Additional Neer categories involve fractures associated with an anterior or posterior humeral head dislocation and fractures involving the articular surface of the humeral head. *Appendix 2* shows the Neer classification system, with numbering of the categories by Sidor *et al.*¹⁹

The dotted lines in *Figure 1* represent lines of epiphyseal scar where the four segments that ossify separately fuse at skeletal maturity. Codman made a key observation that fractures tend to generally occur along lines of these epiphyseal scars in the proximal humerus.²⁰

A large prospective epidemiological study involving 1027 fractures found that around half (51%) of proximal humeral fractures are displaced according to Neer's criteria.⁷ Of these, the largest groups involved two-part surgical neck fractures (28% of the whole population) and three-part greater tuberosity and surgical neck fractures (9%). Three-part lesser tuberosity and surgical neck fractures were rare (0.3%). Four-part fractures without fracture dislocation comprised 2% of the total. Overall, approximately 40% of proximal humeral fractures involve a displaced surgical neck. *Figure 2* shows an example of a two-part fracture of the surgical neck and subsequent internal fixation.

Limitations of the Neer classification include the arbitrary thresholds for displacement; the absence of key potentially prognostic fracture patterns, particularly those relating to the relative positions of the shaft and the humeral head (varus and valgus displacement); and the findings of only fair to moderate interobserver and intraobserver agreement for classification.^{21,22}

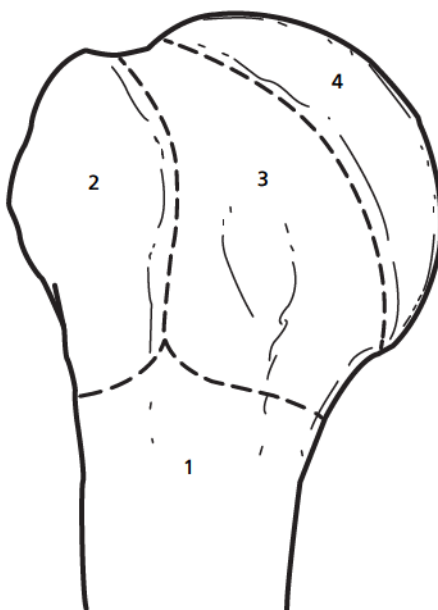


FIGURE 1 Proximal humerus anatomy and the four segments. (1) Shaft, (2) greater tuberosity, (3) lesser tuberosity and (4) head.

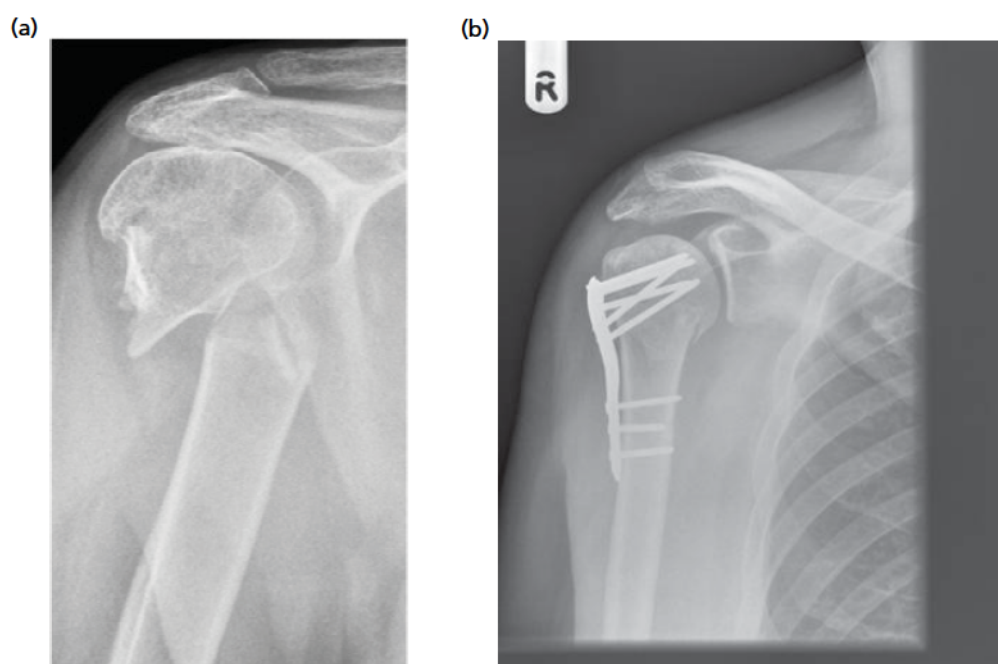


FIGURE 2 Two-part fracture of the (a) surgical neck with (b) subsequent internal fixation with plate and screws.

Treatment: care pathways

People with these injuries typically present with a very painful and swollen shoulder to an accident and emergency (A&E) department soon after their injury. As well as having severely limited movement of the shoulder, there is often deformity and bruising. On radiographic (generally plain radiography) confirmation of their fracture, patients are either sent home with their arm in a sling and given analgesics or admitted to hospital for social or other health-related reasons. (In rare instances there are serious complications, such as an open fracture or major vascular injury, necessitating emergency surgery.) Subsequently, patients are seen by an orthopaedic specialist in the next fracture clinic at the hospital. A second set of radiographs may be taken to further characterise the fracture. These are likely to be at least two views from the standard shoulder trauma series: anteroposterior in the scapular plane, scapular Y-lateral and axillary or modified axillary.^{23,24}

The majority of these fractures, including all undisplaced fractures, are treated non-surgically. Closed reduction of displaced fractures to reposition the fractured parts is generally not carried out. Non-surgical treatment generally involves 'immobilisation' and support of the injured arm in a sling for around 3 weeks²⁵ to facilitate initial fracture healing before allowing shoulder movement. Surgical treatment, which is often undertaken for more complex displaced fractures, usually involves either closed or open reduction of the fracture and internal fixation using various devices or replacing the humeral head with a hemiarthroplasty. In some cases, total joint replacement (including reverse polarity shoulder arthroplasty in which the ball part is uppermost, fixed to the socket, and the cup is fixed to the upper end of the humerus) may be performed. Common methods of internal fixation are plating, in which a metal plate is placed alongside the repositioned bone and secured using screws (see *Figure 2b*); intramedullary nailing, in which a nail is inserted into and along the medullary canal and usually 'locked' into place, such as with screws; and percutaneous fixation using pins, screws and/or wires after closed reduction. Bone grafts or substitutes may be used to fill bony voids. Typically, patients are placed on the next available list of a specialist surgeon and operated on within 3 weeks of their injury. Post-operative treatment generally involves a period of immobilisation. Whether treated surgically or non-surgically, rehabilitation, comprising a mixture of advice, instruction, physiotherapy and home exercises, is required, with the aim of restoring shoulder function and achieving functional independence.²⁵

Evidence underpinning practice

Historically, there has been a paucity of evidence from randomised controlled trials (RCTs) to inform the management of these fractures. In 2007, at the time of developing the PROximal Fracture of the Humerus: Evaluation by Randomisation (ProFHER) trial protocol, the Cochrane review (search date September 2006) on this topic included just 12 small trials (578 participants).²⁶ Of these, three heterogeneous and methodologically flawed trials (102 participants in total) compared surgery with non-surgical treatment.^{27–29} The review concluded that there was ‘insufficient evidence from RCTs to determine which interventions are the most appropriate for the management of different types of proximal humeral fractures’.²⁶ More specifically, it concluded that ‘It is unclear whether operative intervention, even for specific fracture types, will produce consistently better long term outcomes’.²⁶ The review noted that difficulties in patient recruitment seemed a key problem in this area, particularly in trials involving surgery, as shown by the abandonment of three trials and the slow rates of recruitment and lower than planned recruitment rates in other trials.

Subsequent updates of the Cochrane review have noted a continuing inadequacy of the evidence. The update published in 2012 (search date January 2012) included six trials (270 participants) comparing surgery with non-surgical treatment but noted also six ongoing trials that aim to recruit 1052 patients in all.³⁰ All three of the more recent trials^{31–33} recorded validated patient-reported outcome measures of upper limb function. Such measures, which include the Disabilities of the Arm, Shoulder and Hand questionnaire³⁴ and the Oxford Shoulder Score (OSS),³⁵ were unavailable previously.

As well as a lack of evidence to inform on all aspects of management of these fractures (including optimal conservative or surgical treatment), there is evidence of substantial variation in practice. This includes the length of immobilisation and the timing and provision of physiotherapy for less severe fractures²⁵ and the use of surgery.^{15,36,37} Additionally, technology is changing all the time and there are various pressures for early implementation of new fixation devices^{15,37} and surgical techniques, for example.³⁸

Rationale for a trial comparing surgical with non-surgical treatment

Surgery is rarely essential for these fractures and there is a clinical consensus that it is not needed for the majority of non-displaced or minimally displaced fractures. As shown by the large variation in practice, there is uncertainty about the role of surgery for the majority of displaced fractures, most of which involve the surgical neck. The commonly perceived advantages of surgery are restoration of anatomy, the potential avoidance of symptomatic nonunion and malunion, and improved functional outcome. In particular, this last presumption is unproven.³⁰ Moreover, surgery involves additional trauma and risks in terms of anaesthesia and complications (such as wound infection) and sometimes surgical errors.³⁹ There is a need also to justify the greater initial costs associated with surgery, including the potentially increased inpatient stay. Later surgery for failed non-surgical treatment or major revision surgery for surgical complications is usually technically more difficult and less successful. There is some limited evidence indicating that initial surgery is associated with a higher risk of subsequent surgery.^{15,30}

Aim of the ProFHER trial

This key treatment uncertainty formed the rationale for the ProFHER trial. This was a pragmatic multicentre RCT whose primary objective was to evaluate the clinical effectiveness and cost-effectiveness of surgical compared with non-surgical treatment of the majority of displaced fractures of the proximal humerus involving the surgical neck in adults.

This objective was supported and supplemented by the following secondary objectives:

- To describe the study population in terms of key characteristics and in particular to categorise, using optimised methods including training, the trial fractures according to the Neer classification through an independent assessment of baseline radiographs of all randomised patients in the ProFHER trial.
- To ensure that both groups of patients receive comparable and good standards of care. This included remedying identified 'gaps' such as through the development of a leaflet to advise patients on self-care during sling immobilisation and a physiotherapy protocol in consultation with specialist shoulder physiotherapists and promoting the need to encourage home exercise.
- To explore the effect of age (< 65 years vs. ≥ 65 years), type of fracture (involving none vs. one or both tuberosities) and patients' treatment preference on the primary outcome (OSS).
- To explore surgeon treatment preferences, including the relationships between surgeons' decisions to exclude patients because of a lack of equipoise, and patient age (< 65 years vs. ≥ 65 years) and fracture type (involving none or one or both tuberosities).

Chapter 2 Methods

Summary of trial design

The ProFHER trial was a pragmatic UK-based multicentre RCT, with a concomitant economic evaluation, of surgical compared with non-surgical treatment of acute displaced proximal humeral fractures involving the surgical neck in adults.⁴⁰ The choice of surgical intervention was the decision of the surgeons, who had to use techniques with which they were experienced and fully familiar. Good standards of non-surgical care, including care-provider competence, and a comparable provision of rehabilitation for both groups of patients were expected. The latter was facilitated through the provision to patients of an information leaflet on personal care during initial sling immobilisation and the development of a written physiotherapy protocol. We aimed to recruit and randomise 250 patients with these fractures and follow them up for 2 years. Our primary outcome was the OSS,^{35,41} collected at 6, 12 and 24 months via a postal questionnaire. There was no blinding of outcome assessment. However, coding, including the independent categorisation of baseline fractures, was carried out blind to allocation. All analyses were based on intention to treat (ITT). The final version of the study protocol (version 7), shown in *Appendix 3*, was published in 2009.⁴⁰ *Appendix 4* summarises the changes to the protocol.

Randomisation: sequence generation and allocation concealment

Patients were randomised on an equal basis to surgery or non-surgical treatment by the remote randomisation service (telephone or web based) provided by the York Trials Unit (YTU), University of York. Throughout, randomisation was stratified by tuberosity involvement (yes or no) with allocation using random block sizes (these were four, eight and 12). The original intention, after a specified time, to use minimisation based on the minimisation factors of tuberosity involvement (yes or no) and study centre was not implemented (see *Appendix 4*). These measures were taken to avoid the risk of prediction of allocation. As fewer patients than expected were being recruited at centres, the initial concern about an imbalance leading to an excess in the allocation to surgery at a centre, with the associated cost and logistics of performing this surgery, no longer applied. Allocation concealment was assured by the treatment being assigned only after obtaining patient identifiers and key baseline data.

Participant eligibility

Only adults with an acute displaced unilateral fracture of the proximal humerus involving the surgical neck were eligible for inclusion. Allowance was made for surgical neck displacement that did not meet the exact displacement criteria of the Neer classification (1 cm and/or 45° angulation of displaced parts) when this reflected an individual surgeon's equipoise (e.g. whether or not the surgical neck fracture was displaced enough to be considered for surgical treatment). Fractures involving dislocation of the humeral head out of the glenohumeral joint were explicitly excluded. Although fractures involving the articular surface (head-splitting fractures) were not explicitly excluded, it was anticipated that these rare fractures would not be considered eligible.

The following inclusion and exclusion criteria, which were applied throughout patient recruitment, were summarised on the study eligibility form (see *Appendix 5*). (The order of the exclusion criteria was adjusted for presentational reasons.)

Inclusion criteria

Adults (aged ≥ 16 years) presenting to the participating hospital within 3 weeks of their injury with a radiographically confirmed displaced fracture of the proximal humerus involving the surgical neck were eligible for inclusion. This included all two-part surgical neck fractures and three-part (including surgical neck) and four-part fractures of the proximal humerus (according to the Neer classification). Also included were displaced surgical neck fractures that did not meet the exact displacement criteria of the Neer classification.

Exclusion criteria

- Associated dislocation of the injured shoulder joint.
- Open fracture.
- Lack of mental capacity: unable to understand the trial procedure or instructions for rehabilitation; significant mental impairment that would preclude compliance with rehabilitation and treatment advice.
- Comorbidities precluding surgery/anaesthesia.
- A clear indication for surgery such as severe soft-tissue compromise requiring surgery/emergency treatment (nerve injury/dysfunction).
- Multiple injuries: same limb fractures; other upper limb fractures.
- Pathological fractures (other than osteoporotic) and terminal illness.
- Participant not resident in the hospital catchment area.

Also listed on the study eligibility form was 'Other reason to exclude the patient', with an open text box to record the reason.

Sample population

All patients meeting the main inclusion criteria (age ≥ 16 years, presenting within 3 weeks of injury, with a displaced fracture of the proximal humerus involving the surgical neck) were considered potentially eligible. Ineligible patients were defined as those who were excluded for one or more reasons given in the list of exclusion criteria or for another reason, such as a surgeon's preference for one of the treatment options, reflecting lack of surgical equipoise, stated by the surgeon. Some eligible patients were not included because of a lack of patient consent (non-consenting patients), reflecting a strong preference for a specific treatment option or refusal of randomisation.

Study setting

Trial recruitment was undertaken in the orthopaedic departments of 33 of the 35 participating acute NHS hospitals; 33 participating hospitals were based in England and one each was based in Scotland and Wales. Patient care pathways including outpatient and community-based rehabilitation were as per routine practice. Patient recruitment, which was from fracture clinics or orthopaedic wards, lasted from 17 September 2008 until 13 April 2011, when the last patient was recruited. Follow-up was for 2 years.

Interventions

Each centre in the trial had to agree to forgo the introduction of radically novel and experimental interventions for these fractures during the recruitment period. The centres were also responsible for identifying surgeons with sufficient expertise for participation in the trial. We emphasised to the participating centres that central to obtaining reliable evidence was that good standard care, both surgical and non-surgical, was provided throughout the trial. It was expected that, as part of good standard care,

surgery for these types of fractures would usually be carried out by specialist surgeons, who were experienced in operating on these fractures. When possible, decisions about the method of surgery, when allocated, and non-surgical treatment were left to the clinical judgement of the participating clinicians. Participating surgeons were advised that they must, however, use surgical interventions and procedures with which they were familiar (and thus experienced with). This was to avoid learning curve problems.^{42,43} Similarly, other clinicians, including physiotherapists, were advised that they should use procedures with which they were familiar. Additionally, prescribed deviations from the planned surgery and the physiotherapy protocol for rehabilitation provided for the trial had to be documented and reasons given.

Surgery

We anticipated that plate fixation or intramedullary nailing were likely to be chosen for displaced (two-part) surgical neck fractures. For three-part (including displaced surgical neck) or four-part fractures, we anticipated that surgical interventions were likely to include internal fixation such as nails, plates or other methods that preserved the humeral head, or humeral head replacement (hemiarthroplasty).

We stipulated that peri- and post-operative management, including anaesthesia and analgesia, antibiotic and thromboembolism prophylaxis and dressing policies, should follow local guidelines and that similar mobilisation protocols should be provided for all interventions.

Non-surgical intervention (control)

We recommended that conservatively treated patients should be given sling immobilisation for about 3 weeks or for as long as the treating clinician deemed necessary and active early rehabilitation. We stressed the need for competence in conservative methods, including rehabilitation.

Rehabilitation

We stipulated that all trial participants who were assessed as eligible for the trial should be provided with written advice on personal care during sling immobilisation. We stressed that similar access to physiotherapy should be provided for surgical and non-surgical participants. A basic treatment protocol for physiotherapy was provided. This emphasised that, although the protocol acted as a guide, variation in practice was accepted and anticipated. However, electrotherapy, other than transcutaneous electrical nerve stimulation (TENS), was discouraged. (As there is no evidence in support of TENS for treating these fractures,³⁰ its exclusion served to avoid potential confounding given its still albeit unproven effect on fracture healing.⁴⁴) We promoted strongly the need to encourage patients to perform home exercises and that they should receive information sheets illustrating how to carry out the exercises.

Outcome measures

Primary outcome measure: the Oxford Shoulder Score

The primary outcome measure was the OSS,³⁵ assessed at 6, 12 and 24 months. We purposefully did not stipulate a primary time point as we considered that none of these three was of primary interest; thus, the primary analysis compared the results of the two groups over all three follow-up assessments. The OSS is a condition-specific questionnaire providing a total score based on the patient's subjective assessment of pain and impairment of activities of daily living (ADL). The OSS contains 12 items, each with five categories of response. It has been shown to correlate well with both the professionally endorsed Constant–Murley shoulder score⁴⁵ and the 36-item Short Form health survey (SF-36)⁴⁶ and to be sensitive to clinical change at 6 months after surgical intervention.⁴⁷ It has been demonstrated to be consistent, reproducible and valid in a UK population.³⁵ Based on the revised scoring algorithm used in our analyses, the range of available scores was from 0 (worst) to 48 (best).⁴¹ OSS data were collected by participant-completed questionnaires that were posted to participants at 6, 12 and 24 months. OSS data were not collected at baseline because questions refer to functioning during the past month and baseline assessments usually took place within a week of injury.

Secondary outcomes

The following secondary outcomes were collected:

- health status/patient-reported quality of life data measured by:
 - the 12-item Short Form health survey (SF-12)⁴⁸ at 6, 12 and 24 months
 - the European Quality of Life-5 Dimensions (EQ-5D)⁴⁹ at 3, 6, 12 and 24 months (also collected at baseline)
- surgical and other shoulder fracture-related complications
- secondary surgery to the shoulder or increased/new shoulder-related therapy
- medical complications during the inpatient stay
- mortality.

12-item Short Form health survey

The SF-12 is a generic health status measure and a short form of the SF-36 health survey. It consists of 12 questions measuring eight domains (physical, role physical, bodily pain, general health, vitality, social functioning, role emotional and mental health) rated over the past month. Questions have three or five response categories and responses are summarised into a physical component summary (PCS) score and a mental component summary (MCS) score. Outcomes range from 0 (lowest level of health) to 100 (highest level of health). SF-12 data were collected by patient-completed questionnaires at 6, 12 and 24 months. (The SF-12 PCS and MCS scores are presented here.)

European Quality of Life-5 Dimensions

The EQ-5D is a standardised measure of current health status developed by the EuroQol Group for clinical and economic appraisals. We used the EQ-5D-3L version. This consists of five questions assessing five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression). Each dimension is rated on three levels: no problems, some problems and extreme problems. Scores are converted to a single summary index ranging from 0 (death) to 1 (perfect health), although a score of < 0 (worse than death) is possible too. EQ-5D data were collected by patient-completed questionnaires at baseline and 3, 6, 12 and 24 months. (The EQ-5D data are analysed as part of the economic evaluation and are not presented separately in clinical effectiveness analysis.)

Surgical and other shoulder fracture-related complications, secondary surgery to the shoulder and increased/new shoulder-related therapy

These outcomes were extracted from hospital records by hospital staff, and centres were required to complete relevant case report forms (CRFs) at the end of the inpatient stay, when applicable, and again at 1 and 2 years' follow-up. Listed categories were surgical site infection, haematoma formation at the surgical site, nerve injury, implant-related complication/failure, dislocation/instability, metalwork problems, avascular necrosis (AVN), nonunion/malunion and 'other'.

Medical complications during the inpatient stay

Early medical complications were documented by centre staff on a CRF for recording episodes during the inpatient stay. Listed categories were myocardial infarction (MI), stroke, chest infection, other hospital-acquired infection, deep-vein thrombosis (DVT), pulmonary embolism, admission to the intensive care unit, admission to the high-dependency unit and other serious event.

Mortality

Mortality at any time was reported by centre staff to the YTU using a CRF.

Data collection schedule

An overview of the time points at which trial data were collected is presented in *Table 1*.

TABLE 1 Data collection schedule

Outcome	Baseline	End of inpatient episode	3 months	6 months	12 months	24 months
Demographics and fracture data	X					
OSS (primary outcome)				X	X	X
SF-12				X	X	X
EQ-5D	X		X ^a	X	X	X
Surgical and other shoulder fracture-related complications		X			X	X
Medical complications		X				
Secondary surgery to the shoulder		X			X	X
Increased/new therapy for shoulder-related complications		X			X	X

a Collection of data at 3 months was restricted purposefully to those required for the economic analysis. Six months was considered the earliest time point that was appropriate to collect OSS and SF-12 data in terms of the typical recovery pattern and treatment regimens of patients with these fractures. It was expected that many patients would still be undergoing rehabilitation, primarily physiotherapy, at 3 months.

Data for the economic evaluation

Health utility data were obtained from the EQ-5D data collected at baseline and 3, 6, 12 and 24 months from patient questionnaires. Prospective cost data for trial participants were collected from the hospital forms and patient questionnaires. These included costs incurred in hospital (surgery, inpatient stay and physiotherapy) and subsequently [outpatient physiotherapy, visits to a general practitioner (GP) and other NHS costs]. We collected surgical data on the actual procedures carried out, including anaesthesia, and the interventions provided and the experience of operators/care providers (according to grade).

Data were collected to allow some estimation of other 'societal' costs to the patients, specifically private health care, days off paid work or days unable to perform normal activities; but, as indicated in the protocol, these data were not used in the analyses presented in this report.

Classification of study fractures

Copies of all baseline radiographs for all randomised patients were collected. These were processed to facilitate an independent and blinded classification of the study fractures based on the Neer classification system at the end of study recruitment by two orthopaedic specialist shoulder surgeons experienced in assessing and operating on proximal humeral fractures. This was preceded by a training session on the Neer classification, which has been shown to improve inter-rater agreement.²²

Sample size

The primary outcome for the trial is differences in patients' subjective assessments of pain and ADL as measured by the OSS. To justify both the increased cost of surgery and the exposure to the hazards of surgery, there need to be greater improvements in patients' subjective assessments of pain and ADL in the surgery group than in the conservative treatment group. There continues to be a lack of data to inform on the minimum clinically important difference for any validated patient-reported outcome measure for these fractures. Observational data collected by the chief investigator and his colleagues found that those patients who had surgery without subsequent complications [1- to 5-year data from 50 patients treated with a PHILOS® plate (Synthes Ltd, Welwyn Garden City, UK) between 2001 and 2005] had a 5-point differential

improvement in OSS compared with patients treated conservatively (3- to 12-month data from 103 patients treated between April 2000 and July 2003).⁵⁰ This 5-point difference was judged to be of clinical importance. Given a standard deviation (SD) of 12 this equated to an effect size of 0.42. It was calculated that to detect a minimum effect size of 0.4 with 80% power and a two-sided 5% significance level 100 patients in each treatment group would be required. After allowing for a potential loss to follow-up of 20%, we proposed to recruit and randomise 250 patients (125 in the surgery group and 125 in the control group). (This did not take account of any potential cluster effect, such as from similarities between surgeons within centres, as we did not expect there to be many patients treated by individual surgeons.⁵¹)

Estimating the numbers of patients and centres needed to recruit

There were insufficient good-quality prospectively collected data to inform our estimates of the number of trial centres and duration of recruitment needed to achieve our target of 250 trial participants. Although data from the epidemiological study conducted in Edinburgh⁷ were used, particularly in terms of the relative proportions of Neer categories in our intended fracture population, using these as a basis for estimating the sample populations in individual hospitals was hampered by the unavailability of accurate data on the catchment populations of individual hospitals. There is also the unknown but expected variation in risk factors in these populations. Another approach of directly asking putative principal investigators when developing the study proposal yielded rough estimates without reliable data on how many proximal humeral fractures were treated at a centre and what proportion of these might be treated surgically. In December 2006 we sent each of the six putative principal investigators a form requesting information that would inform on the recruitment potential of the site. This included estimates of the numbers of people presenting with proximal humeral fractures at individual centres each year, how many of these had surgery and how many would be eligible for inclusion. We received two completed forms, one provided by the chief investigator. This lack of information on centre activity relating to proximal humeral fractures was mirrored throughout the trial. Furthermore, we used HES data to obtain a general impression of overall activity at a hospital (taken as a proxy for catchment population) and specific proximal humeral fracture activity, while recognising that patients with less serious (minimally displaced) fractures may not have been admitted. *Appendix 6* gives details of the calculations used to estimate the size of our population sample using the HES data.

Proposed recruitment rate

From the outset we anticipated that recruitment to this trial would be challenging and thus we set very conservative recruitment targets. We set our recruitment period at 18 months and aimed to recruit between 18–20 centres. We estimated that across 18 centres there would be 6132 patients (corrected from the protocol) with a proximal humeral fracture over the 18 months of recruitment. Based on data from Court-Brown *et al.*,⁷ we calculated that 2391 patients would have the fracture types suitable for inclusion in the trial. Thus, to achieve our sample size we needed to recruit only 11% of these patients. We established an expectation that each participating centre would recruit at least one participant per month.

Because of the uncertainty in terms of recruitment we proposed a feasibility study in which there was an assessment point after 10 months of recruitment to determine if recruitment into the trial was sufficient (set at 88 patients) to justify its continuation after 18 months. Subsequently, we successfully argued for the cancellation of this and the associated stopping rule at 12 months (see *Appendix 4*).

Governance

Ethical arrangements

Multicentre Research Ethics Committee (MREC) (now the National Research Ethics Service Committee Yorkshire and the Humber – Leeds West) approval was obtained from the York Research Ethics Committee (reference number 08/H1311/12; date of favourable ethical opinion 11 March 2008). This was preceded by approval from the School of Health and Social Care Research Governance and Ethics Committee, Teesside University (sponsor) via the Chair's action (date of favourable ethical opinion 10 December 2007).

Separate approval was also sought from local research ethics committees for each centre up until this became the responsibility of NHS research and development (R&D) offices from 1 April 2009. Approval was obtained from all R&D offices of participating hospitals. In some cases R&D approval was required also for primary care trusts to allow for data collection by physiotherapists delivering care in the community such as in general practices or other local hospitals.

Appendix 7 lists the six minor and seven substantial amendments to the protocol approved by the MREC. The R&D office of each participating hospital was notified of all amendments after approval was received from the MREC. The last three amendments, dating from September 2010, apply to the extension to follow-up to enable patient questionnaires to be collected at 3, 4 and 5 years. The long-term follow-up study will be reported separately in 2016.

The MREC was notified within 15 days of the YTH being notified of any serious adverse events (SAEs) that were judged to be both unexpected and related to trial participation. Reports to the MREC were submitted on an annual basis.

Trial governance

Independent supervision

Independent and non-executive members of the Trial Steering Committee (TSC), including the independent chairperson, two other independent members (a shoulder specialist and a statistician), two service users and the sponsor representative provided independent supervision and advice on the trial. Up to the final TSC meeting, whole-population data on accrual, recruitment and follow-up, as well as 'housekeeping' statistics, were disseminated to TSC members on a regular basis. The TSC, which convened at the discretion of the independent chairperson, met formally with members of the extended Trial Management Group (TMG) a total of seven times.

A separate Data Monitoring and Ethics Committee (DMEC), comprising an independent statistician, a shoulder specialist and a consultant rheumatologist, met a total of eight times, generally via audioconference. An analysis of major complications was provided to the DMEC on a regular basis following the commencement of trial recruitment.

Appendix 8 lists the membership of the TMG and these two committees.

Good Clinical Practice

Throughout the trial we adhered to the good clinical research practice guidelines.⁵² Research governance approval was obtained from participating trusts. This included local modifications of the Clinical Trial Agreement (or contract) between the trial sponsor and local trusts and agreeing on which hospital staff needed Good Clinical Practice (GCP) training.

Patient consent and information

Written informed consent was obtained from all participants. A participant information sheet for the trial, which was produced with the involvement of service users, was provided to all potential participants (see *Appendix 9*). This gave a balanced and easily understandable account of the possible benefits and known risks of the interventions under test and stated explicitly that quality of care would not be compromised if the participant decided to (1) not enter the trial or (2) withdraw their consent. We allowed patients sufficient time to make an informed decision without unduly delaying their care. (Although we intended, when possible, to provide translations of the participant information sheet for non-English speakers, none was requested throughout the trial.)

Blinding and data management

There was no blinding of trial participants, care providers or outcome assessments. Data entry, processing and analysis were carried out independently. Aside from accrual and whole-population baseline statistics, interim results were not made available to the trial investigators, including the chief investigator, or associates in the participant centres. All data, including baseline radiographs, were anonymised before distribution using study IDs, which were unique to each participant. Coding, which always involved at least two independent coders, was carried out blind to allocation. To enable appropriate data checking, the trial statistician conducting the analyses was unblinded to treatment group. However, the primary analyses were repeated by a second blinded statistician using different statistical software and the results were in agreement.

The anonymity of all treatment providers, including surgeons and physiotherapists, and centres is preserved for all analyses.

Retention of all relevant trial documentation is set for a minimum of 20 years from the date of trial completion.

Data handling

The patient questionnaires and hospital forms (i.e. CRFs) were designed using TeleForm software (version 10; Cardiff Software, Cambridge, UK). The CRFs were marked up in TeleForm with variable names and appropriate scoring. When the CRFs were completed and returned to the YTU they were prepared for scanning by a data entry clerk using the TeleForm software. When a form would not scan, the data were manually entered. When a form was scanned, the data were then verified depending on what TeleForm had identified as requiring correction. The verified data were then downloaded into the study database and were available for second checking. This involved each hard copy of the CRF being compared against the entry stored in the study database and correcting the electronic data as necessary. All data that had been scanned, downloaded and second checked were then ready for validating. The data validation was undertaken by the data manager who applied predetermined rules to check that the data recorded for the variables on the CRF, such as whether or not the dates when the patient injured his or her shoulder and was assessed for eligibility fell within the 3-week time frame for being eligible for the trial. Data that had been validated were then available to the trial statistician and health economist.

Data checks, cleaning and processing

The MREC did not allow any return to patients for missing data. Data checks, including those conducted by the data manager, alerted us to recurrent problems with completion of hospital forms during trial recruitment.

Data 'cleaning', before the staged closure of the data set according to pre-established time points, was carried out in a systematic manner with consultation on and discussion of any amendments.

Categorisation and interpretation of text comments that were recorded on the various data collection forms was generally performed by two independent and blinded members of the TMG, who then discussed disagreements to reach a consensus. The criteria for assessment were documented, often beforehand. When categorisation was performed by one rater, a second rater was asked to check the results and criteria.

Clinical effectiveness analysis

All analyses pertaining to clinical effectiveness were conducted to a prespecified analysis plan, endorsed by the DMEC. Signatures of approval of the final version (version 19; August 2013) were obtained from the chief investigator (AR), trial manager (SB) and 'co-applicant' (HH) before release of results by the statistician (AK). ITT analysis, including all randomised patients in the groups to which they were randomised, was conducted throughout. All main analyses were performed using Stata 12 (StataCorp LP, College Station, TX, USA) using two-sided significance tests at the 5% significance level.

Primary outcome analyses

The primary outcome (OSS) was summarised descriptively by treatment group at each time point (6, 12 and 24 months post randomisation), including the extent of missing data. Baseline characteristics predicting non-response were determined by logistic regression. As no specific assessment time was considered to be of primary interest, the primary analysis compared the surgical and non-surgical treatment groups over all follow-up assessments. A random-slope multilevel model was fitted to the data with time points nested in patients to allow for clustering of data within each patient. This model adjusted for the fixed effects of time (6, 12 or 24 months), tuberosity involvement at baseline (yes or no), age (< 65 years or ≥ 65 years), gender, health status at baseline (EQ-5D), treatment group and the interaction between treatment and time (to assess whether or not any difference between treatment groups changed over time). Different covariance patterns for the repeated measurements were explored and the most appropriate pattern was used for the final model. Estimates of the difference in OSS between treatment groups were assessed overall and at individual time points (significance set at $p < 0.05$). The trial was powered to detect an effect size of approximately 5 points.

Subgroup and sensitivity analyses

Two subgroup analyses, the second with an associated sensitivity analysis, were set up a priori; these were for the primary outcome only (OSS):

- age (≥ 65 years vs. < 65 years)⁵³
- surgical neck fractures involving one or both tuberosities (no vs. yes) as per initial surgeon classification (Neer one- and two-part fractures vs. Neer three- and four-part fractures based on classification by two independent experts from radiographs was used as a sensitivity analysis).

To strengthen our choice of subgroups, we made the following a priori specification of subgroup direction:⁵⁴

- the benefit of surgery over non-surgical treatment will be larger in patients aged < 65 years
- the benefit of surgery over non-surgical treatment will be larger in patients with a surgical neck fracture with displacement of one or both tuberosities (Neer three- and four-part fractures).

The primary analysis model was extended for each of these subgroup analyses by including in turn an interaction of treatment group with age group, tuberosity involvement or Neer parts. The differences in model fit between these and the base model were assessed at $p < 0.10$. The statistical significance level of 10% was chosen as the study had not been powered to detect interactions.

Smoking status was added to the primary model as a sensitivity analysis (see *Appendix 4*). To ascertain the impact of prior beliefs, a treatment group \times patient treatment preference interaction was added to the base model. The variability in OSS across recruitment sites was analysed descriptively.

Secondary outcomes

The SF-12 PCS and MCS scores were analysed by multilevel modelling using the same predictors as for the OSS primary model.

The numbers of surgical and shoulder fracture-related complications as well as treatments for these were reported by treatment group and time point. The frequencies of shoulder surgery and fracture-related complications and any treatments for these were compared between treatment groups using the chi-squared test. Mortality rates were reported and compared between treatment groups using the chi-squared test. The use of the chi-squared test was determined by prespecified criteria.

Other outcomes

The numbers of other admissions and treatments for serious newly diagnosed medical complications and fractures reported over the 2-year follow-up period were listed by treatment arm. Selected baseline characteristics were inspected for patients with or without these complications. Adverse events were summarised for each arm by type and expectedness/relatedness to treatment. Complications reported within the physiotherapy treatment logs were reported descriptively if they were identified via the inpatient episode or hospital follow-up forms.

Lack of surgeon equipoise

A separate analysis of lack of equipoise across centres and individual surgeons was undertaken. A lack of equipoise occurred when the sole reason why a patient was found to be ineligible was because the surgeon's response to the question about 'other reason to exclude the patient' when completing the study eligibility form was that he or she had a treatment preference. To determine whether or not the reason for exclusion was the result of a lack of equipoise, the reasons given were checked by two independent raters (AR and HH), who resolved any disagreements through discussion. The impact of lack of surgeon equipoise on study generalisability was explored by comparing participants' age, gender, time since injury and tuberosity involvement between those ineligible because of a lack of equipoise and all eligible participants using a t -test or chi-squared test depending on the type of outcome. The relationship between age and fracture type (the prespecified subgroup analyses) and surgeon treatment preference was explored for those participants excluded because of a lack of equipoise. Age and fracture type were compared between the surgeon treatment preference groups using a chi-squared test.

The Neer classification of baseline radiographs

Baseline fracture radiographs were evaluated during eligibility assessment and used later for an independent classification using Neer's proforma. A separate analysis explored the quality of available baseline fracture radiographs, the agreement between raters and the agreement between the two assessments in terms of the involvement/displacement of relevant anatomical structures. The kappa statistic was used when appropriate.

Economic analysis

Economic analysis was likewise conducted to a prespecified analysis plan shared with the DMEC. Key differences from the protocol are described in *Appendix 4*. The analysis was carried out from the perspective of the UK NHS and Personal Social Services.⁵⁵ Data on health utilities, obtained from the EQ-5D data collected from patient questionnaires, were converted into quality-adjusted life-years (QALYs) for each patient using the area under the curve (AUC) method. Costs were expressed in UK pounds sterling at a 2012 price base.

Differences in mean costs and QALYs at 2 years were used to derive an estimate of the cost-effectiveness of surgery and non-surgical treatment. Multiple imputation^{56,57} instead of the method of Lin *et al.*⁵⁸ was used to derive the data set for the base-case analysis. ITT analysis was conducted throughout. An additional analysis was conducted for complete cases (in which patients with any missing data are excluded). All analysis and modelling were undertaken in Stata 12. The full methods used are detailed in *Chapter 7*.

Chapter 3 Further methods: standardisation of trial procedures and care programmes, patient recruitment and data collection and processing

This chapter gives a detailed account of various aspects of trial design and conduct that collectively contributed to the successful meeting of the primary objective of the trial. It begins by describing specific measures taken to ensure standardisation of trial procedures and care programmes. These include processes for setting up trial centres for recruitment and actions taken to meet our secondary objective of ensuring that both groups of patients received comparable and good standards of care. It then describes typical processes involved in patient recruitment and systematic approaches undertaken for data collection and processing. It ends by covering the methods undertaken to meet another secondary objective, that of obtaining a reliable characterisation of all of the fractures included in the trial according to the Neer classification.

Standardisation of trial procedures and care programmes

Throughout the conduct of the trial we emphasised the importance of good practice from clinical and research perspectives; standardised protocols and care pathways; and comparable and sufficient expertise of care providers in the two groups. As far as possible we adopted procedures that reflected and were compatible with usual practice and that did not delay treatment, add unnecessarily to the clinical workload or place a burden on participants. Specific areas of focus are described in the following sections.

Site visits and information

Where possible, a site visit was arranged before submission for local R&D approval of the study to meet all of the key people who would be involved in conducting the trial at that site. In addition to discussion of the rationale and design of the trial, this helped to establish the practical arrangements of trial management at that site and facilitate obtaining approvals for these at the relevant departments (e.g. orthopaedics, physiotherapy and radiology). At this stage or subsequently, the designated person, most often a research nurse, was identified and agreements were reached on which hospital staff needed Good Clinical Practice training and the required modifications to the Clinical Trial Agreement (or contract) between the trust and the trial sponsor. The Clinical Trial Agreement made clear the obligations of the trust to the sponsor and vice versa. The latter included reimbursement according to a payment schedule for data collection by the hospital (see *Appendix 10*). Because several months often elapsed between this site visit and the submission for local R&D approval of the study and subsequent approval, we also undertook a second site visit after R&D approval was obtained to meet with the key people involved in the study. The purpose of this visit was to ensure that all of the practical arrangements were established to enable recruitment to begin on an agreed start date. The site was also provided with its investigator site file, which included all of the essential documentation required to conduct the study and the packs of materials to begin screening and recruiting patients into the trial.

The trial manual, which was developed from the trial protocol to facilitate the conduct of the trial at the hospital sites, reinforced the information on trial governance, patient recruitment, data collection procedures and expectations about trial conduct. It contained copies of all data collection forms, the protocol, information sheets, the physiotherapy protocol and the Clinical Trial Agreement. The designated person completed a staff delegation log, which acted as a record of who was involved in the trial at that site, and a participant log, for documenting actions relating to trial patients. At all times we ensured that consistent messages were given on eligibility criteria and trial procedures.

Radiographs

Initially, we strongly recommended the use of the full shoulder trauma series (three perpendicular views) for assessing fracture eligibility.⁵⁹ Documentation including a PowerPoint (Microsoft Corporation, Redmond, WA, USA) presentation illustrating the full trauma series was made available as part of the trial materials (see *Appendix 11*). We adjusted this expectation after feedback from sites on differences from local practice and anticipated difficulties with implementation. On guidance from independent members of the TSC who emphasised the need to keep a pragmatic approach, we instead promoted the need to adhere to local guidelines for radiographic assessment. We thus stipulated that a minimum of two radiographic views/projections were required for the assessment of study eligibility. Anonymised baseline radiographs were collected for all randomised patients. The radiographic views were recorded on the study eligibility form (see *Appendix 5*).

Assessment of fractures

An introduction to the Neer classification was included in the trial manual that was provided to all hospital sites. In the trial manual and other information for the trial, the study eligibility criteria were expressed in terms of the Neer classification and the displacement criteria stated. However, training in the Neer classification was not provided²² and, in keeping with normal practice in busy A&E departments and fracture clinics, there was no expectation that the recruiting surgeons would classify the fractures. Thus, rather than request that surgeons judge whether or not displaced parts met the Neer criteria, they were asked to indicate on the eligibility form if the fracture *involved* either of the tuberosities (see *Appendix 5*).

Provision of written patient information on personal care during initial sling immobilisation

The initial care of these fractures generally includes immobilisation of the injured arm in a sling, collar and cuff or similar. When working through the processes for patient recruitment in the context of the clinical pathway, we discovered that there was a lack of written patient information on personal care during 'sling immobilisation' at all centres involved in the funding application. This discovery prompted the development of an illustrated information sheet by two physiotherapists, with input and feedback from the TMG and the two service users (both former proximal humeral fracture patients) (see *Appendix 12*). Items covered in the information sheet were the rationale for the sling; sling use and care; advice on keeping mobile; hand and wrist exercises; washing and hygiene; getting comfortable; sleeping position; pain relief, including breathing; important 'don'ts'; warning signs; and a reminder to seek further advice in case of problems. The information sheet was included in the trial recruitment materials to be given to potentially eligible patients. Centres were advised that the ProFHER trial sling immobilisation leaflet could be replaced by a local hospital leaflet if available.

Physiotherapy protocol

The need for equivalent provision of physiotherapy to both treatment groups was emphasised. The absence of a physiotherapy protocol to guide and promote standard care for these fractures was similarly noted at all centres involved in the funding application. With the absence of evidence on which to base practice,²⁶ we aimed to develop a physiotherapy protocol, giving basic treatment guidelines, that was representative of accepted good practice. This process comprised several stages.

Informed by an initial review of current physiotherapy practice at the lead site, a draft protocol structured to reflect successive phases of rehabilitation was drawn up by four specialist physiotherapists based at different sites. The independent TSC advised of the need to seek feedback on the draft protocol and obtain further insights on the current management of these fractures from specialist shoulder physiotherapists and other experts in the field. Additional advice from the TSC pertained to the importance of keeping to routine practice and avoiding unusual interventions, including electrotherapy (other than TENS).

We prepared a short questionnaire (see *Appendix 13*) that asked for feedback on the draft protocol in terms of its acceptability as a guide to basic treatment for these fractures and on whether or not the specialists themselves would adopt the approach given in the protocol. Additionally, the specialists were

asked to list other interventions that they applied routinely for people with these fractures and any interventions available in a NHS setting that they would discourage. In June 2008, the questionnaire and draft protocol were circulated for feedback in two ways: electronically via the hub contacts for the National Physiotherapy Research Network (NPRN)⁶⁰ and in person at the British Elbow and Shoulder Society (BESS) annual scientific meeting in Liverpool.

We received 29 completed questionnaires, predominantly from specialist physiotherapists currently involved in treating these fractures and other shoulder injuries. Fourteen responses were from the 50 allied health professionals attending the BESS meeting, 12 were from the NPRN mailshot (it is unknown how many relevant people received the request for participation) and three were from members of a local shoulder guideline development group.

The results from the questionnaire survey were collated and assessed in terms of the acceptability of the draft protocol and any changes that were needed. Responses to all comments were documented. We found that 27 respondents agreed that the protocol was acceptable as a basis for treatment, and, moreover, were 'happy' with the approach given in the protocol. We made two main modifications to the protocol in response to feedback. The first involved additional text stressing that the protocol was only a guideline and that variation in practice was anticipated and acceptable, with the proviso that any substantive difference was recorded using the hospital proforma. The second was a correction to some contradictory instructions. All of the other interventions listed as being used routinely for people with these fractures were recorded for future reference. None of these interventions was considered unexpected or inappropriate. Notably, there was a strong discouragement of electrotherapy except for the use of TENS for pain relief: 12 of the 14 responses on what interventions should be discouraged specified electrotherapy (including ultrasound), with three excluding TENS from this category.

Subsequent feedback from the TMG and colleagues from the United Kingdom Rotator Cuff Trial on rotator cuff repair⁶¹ on the modified protocol pointed to a greater emphasis on the non-surgical aspects of treatment and rehabilitation and that the protocol applied to both surgically and non-surgically treated patients. The final version of the physiotherapy protocol (see *Appendix 14*) was included in the trial documentation provided to trial sites. The substantial endorsement from other specialist physiotherapists experienced in treating these fractures was stressed in the trial manual.

Home exercises

We promoted the need for physiotherapists to encourage patients to do home exercises and emphasised this in the trial manual and physiotherapy protocol. Home exercises tend to reflect patient ability and to be adapted as the patient recovers, which contradicted the provision of standardised material. Our agreed approach was to check that physiotherapists either provided information leaflets illustrating exercises for home use already or would access a standard web-based facility to generate 'bespoke' exercise sheets. The recording by physiotherapists of whether or not patients had done their home exercises was completed for each session of physiotherapy.

Delivery of physiotherapy

Patients with these fractures often receive physiotherapy at one or more different localities after leaving hospital. This added considerable complexity to the delivery and recording of physiotherapy and outcome. R&D approval for primary care trusts (now disbanded) was sometimes required in addition to that obtained for the acute hospital trusts, to allow for data collection by physiotherapists delivering care in the community, such as in general practices or other local hospitals. For practical purposes, only the main primary care trusts were generally targeted; thus, patients expected to use venues outside their catchment area were ineligible for the trial. Ultimately, additional R&D approval was obtained for 16 primary care trusts.

Contacts for physiotherapy were identified at each trial site to facilitate delivery of physiotherapy and data collection. The contact physiotherapists usually attended the multidisciplinary team meeting at site set-up visits and were named in the local site R&D approval. They were responsible for organising the

physiotherapy including dissemination of trial methods and materials, the provision of advice, the distribution and collection of physiotherapy follow-up forms and often the return of these forms to the YTU. All 32 centres that recruited patients had named contacts for physiotherapy; in 10 centres, the physiotherapy contact was also the designated contact for the centre.

Patient recruitment

This section describes the typical processes involved in the identification of potential participants and trial recruitment. Trial recruitment packs, which included all relevant forms and materials for potentially eligible patients (thus those meeting the trial inclusion criteria), were provided by the YTU. Each potentially eligible patient had a unique four-digit identification number that was prerecorded on the individual forms in the individual recruitment packs. Freepost envelopes were provided. To assist centres further in keeping track of form returns, a participant log listing the various forms and materials for return to the YTU was provided for the designated person to record the dates when individual forms were returned to the YTU.

Identification of potential participants

The designated person at each centre was requested to inform A&E staff about the trial and to distribute specific leaflets about the trial to them. It was important that A&E staff were aware of the trial and also that they took care not to jeopardise patient consent by biasing patients' attitudes to treatment. As per local arrangements, which may have been adjusted specifically for the trial, patients with a radiologically confirmed fracture of the proximal humerus were often referred by A&E to specific fracture clinics. As described earlier, the requirement for a minimum of two perpendicular radiographic views to assess patient eligibility for the trial had been established at site set-up. When feasible, the radiography department flagged up potentially eligible patients for the trial to the orthopaedic department.

In some cases, patients were admitted into hospital from A&E. When this occurred a patient was assessed by the orthopaedic surgeon for eligibility and approached for recruitment into the trial using the same procedures as established for the fracture clinic, as described in the following paragraphs.

At the fracture clinic, the designated person would identify a potentially suitable patient and the orthopaedic surgeon would assess whether or not the inclusion criteria were met as presented in the study eligibility form (see *Appendix 5*). Only when these criteria were met did the surgeon, often with the assistance of the designated person, complete the study eligibility form to confirm whether or not the patient was eligible to take part in the trial (thus, none of the exclusion criteria applied). When the patient was eligible, the surgeon invited the patient to talk to the designated person about obtaining consent to take part in the trial. If the patient was not eligible, the surgeon was required to complete the form by stating what treatment they would advise for the patient. The form was then returned by the designated person to the YTU. For both categories of patient, either the surgeon or the designated person gave the patient the sling immobilisation leaflet supplied in the recruitment pack (see *Appendix 12*) or the hospital equivalent leaflet.

Obtaining informed consent and randomisation

The designated person provided the patient with the patient information leaflet (see *Appendix 9*) and answered their questions before asking for their consent to participate in the trial. Although the patient could decide immediately, it was made clear that he or she could defer the decision for up to 24 hours to think about it further and/or discuss it with family or friends. The result of the decision was recorded by the designated person on the consent status form (see *Appendix 15*). When the patient did not consent, the designated person requested that the orthopaedic surgeon answer the questions on the consent status form on their advised treatment for the patient, the patient's preferred treatment if it had been expressed and the agreed treatment. This form and the study eligibility form were then returned by the designated person to the YTU.

When the patient consented to take part, he or she signed the consent form (see *Appendix 16*) and completed the baseline information form (see *Appendix 17*) before randomisation. Of note is that the baseline information form specifically gave advice about the desirability of stopping smoking for those who admitted they smoked. The designated person then contacted the YTU either by a free telephone service during weekday working hours or via the internet at any time to access a password-protected randomisation service. When randomising the patient, the designated person provided key identification information (e.g. name and date of birth) and basic descriptive details about the patient (gender, fracture tuberosity involvement, location of recruitment: ward or fracture clinic) and answered a series of questions in relation to study eligibility. On receipt of these details and confirmation of study eligibility and patient consent, the random allocation (surgery or not surgery) was given. This was recorded on the baseline information form in addition to informing the patient about the treatment that he or she was to receive.

On randomising an eligible patient, the designated person returned the study eligibility form, consent status form, a copy of the consent form, the baseline information form and copies of the baseline radiographs to the YTU in the Freepost envelope provided. A copy of the consent form was also given to the patient, the original being kept in the medical notes. A letter was sent from the YTU to each trial participant to thank them for agreeing to take part in the trial. Additionally, each patient's GP was notified of the patient's enrolment into the ProFHER trial and provided with a copy of the patient information sheet.

A flow chart showing the patient recruitment and hospital data collection processes is shown in *Figure 3*.

Data collection

Data handling and record keeping

Each trial participant had a unique four-digit identification number, which was prerecorded on all routine data collection forms. A central database at the YTU was developed to prompt the sending out of patient questionnaires and the 1- and 2-year hospital forms, usually by the trial secretary, and to record the return of all forms and facilitate the management of data.

Baseline data

Baseline data were sourced from the study eligibility form, baseline information form and data obtained at randomisation (*Table 2*) and baseline radiographs used for classification of the fracture according to the Neer classification.

Baseline radiographs

In each centre the designated person liaised with the radiography department to arrange for the dispatch to the YTU of copies of the radiographs used in the assessment of trial eligibility for each randomised patient. Before sending the radiographs, either as films or electronically [on compact disc; Joint Photographic Expert Groups (JPEG) files, preferably with a resolution of 300 dpi, were requested], they were identified using a patient's unique four-digit number but were otherwise anonymised, with the removal of patient details such as name.

Patient questionnaires

Postal questionnaires were sent by the YTU to all patients promptly at 3, 6, 12 and 24 months' follow-up after recruitment into the trial. Reply-paid envelopes were included at all times. The questionnaires included the EQ-5D (at all times), the OSS and SF-12 (at 6, 12 and 24 months) and brief questions on the number of consultations with NHS care providers (GPs, physiotherapists, district nurses, etc.), the number of hospital attendances, use of private health care and days lost from work or other normal activities. Copies of the 3-month questionnaire and the 2-year questionnaire are provided in *Appendices 18* and *19* respectively. The content of these questionnaires was commented on by the service users and all versions were approved by the MREC (see, for example, substantial amendment 28 July 2010 in *Appendix 7*). Particular focus revolved around the practicality and ease to the patient of answering questions on health care (care from the NHS,

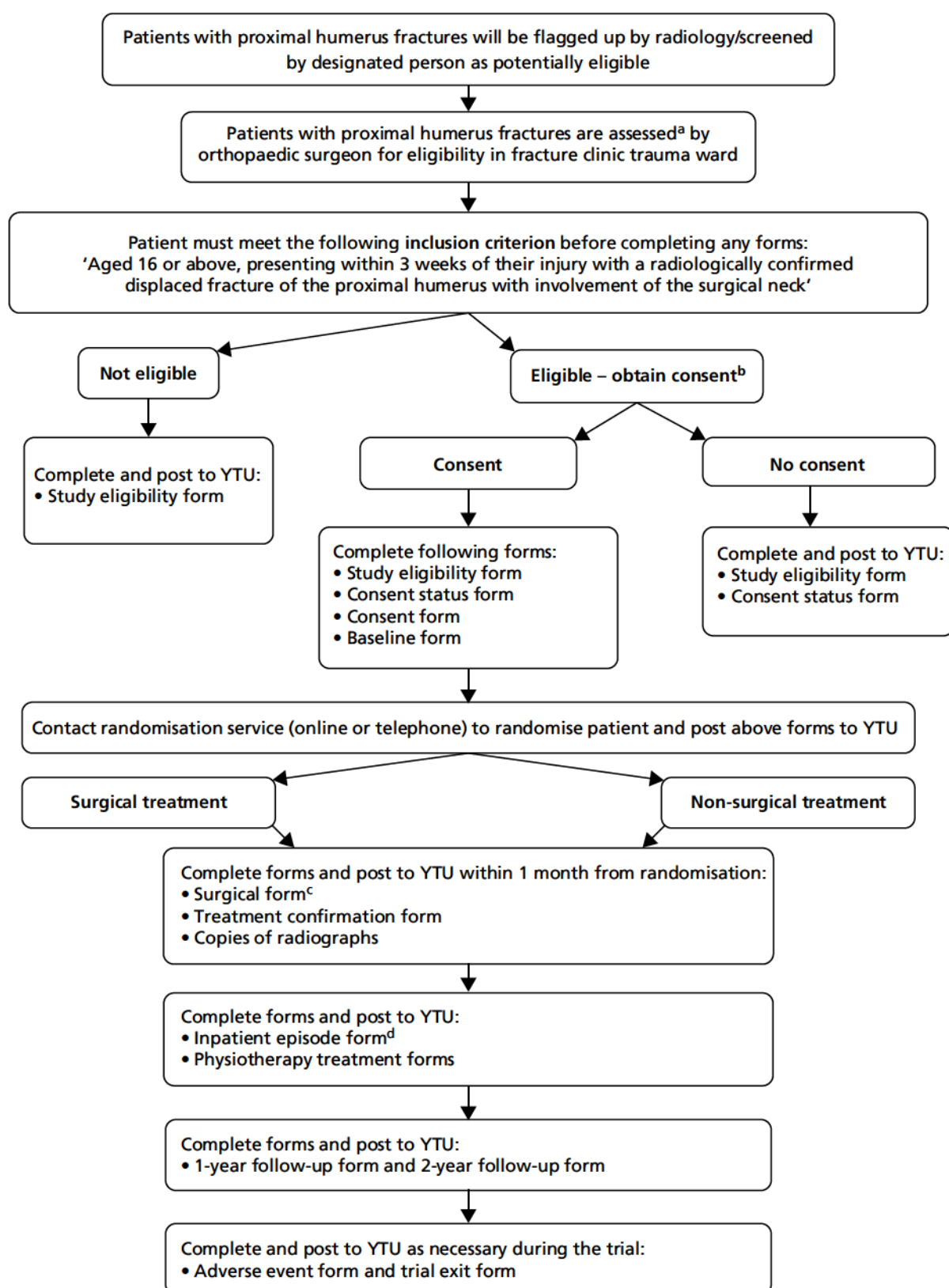


FIGURE 3 Flow chart of patient recruitment and hospital data collection. a, The series of radiographs used to confirm the fracture should be, for example, anteroposterior, axillary (or modified axillary) and scapular Y lateral views; b, after assessing eligibility, patients may need at least 24 hours to decide whether or not to consent to take part; c, the surgical form was also to be completed for any patient randomised to non-surgical treatment who then underwent surgery; and d, the inpatient episode form for a patient not randomised to surgery was not required if the patient was not admitted to hospital.

TABLE 2 Sources of baseline characteristics data

Characteristic	Eligibility form	Baseline form	Randomisation
Gender	X		(X)
Age (date of birth)	X		(X)
Ethnicity		X	
Education (highest qualification)		X	
Employment		X	
Diabetes		X	
Smoking status		X	
Steroid use		X	
Health status (EQ-5D)		X	
Time since injury (date of injury)	X		
Affected shoulder	X		
Tuberosity involvement	X		(X)
Radiographs used	X		
Previous fractures		X	
Previous surgery for fractures		X	
Shoulder dominance		X	
Injury mechanism		X	
Previous treatment preference		(X)	
Location at randomisation (fracture clinic/ward)			X
Treatment allocation		X	X
Date of randomisation			X
(X), data not used for presenting baseline characteristics.			

private treatments) and impact on usual activities. To improve compliance, 2-week and 4-week reminders were sent. Patients were also offered by the trial manager the option of completing a shortened questionnaire (OSS, EQ-5D and information on hospital readmissions) by telephone at 6 weeks. We also sent pre-notification letters at all time points and used unconditional incentives of £5 at 12 and 24 months' follow-up; both approaches are effective at improving response rates.⁶²

The sending out of reminders was recorded as was the timing of questionnaire return. Additionally, patients were asked to add in 'today's date' near the beginning of each questionnaire to record the date when it was completed.

Hospital forms

Seven hospital forms covering various aspects of routine data collection for randomised patients were completed by the designated person or others, as appropriate. These and two other forms for non-routine data collection are described in the following section. *Figure 3* shows the hospital data collection process.

Treatment confirmation form

The (1-month) treatment confirmation form (see *Appendix 20*) was completed at 1 month from the date of randomisation. This multipurpose data collection and 'housekeeping' form was used (1) to establish what treatment the patient received, including prescribed non-surgical treatment, record inpatient

admission and determine the intended or actual start date of physiotherapy and (2) to prompt the dispatch of the baseline radiographs and, when appropriate, the completion and return of the surgical form (see *Appendix 21*) and the inpatient episode form (see *Appendix 22*) to the YTU. The questions on non-surgical treatment and the use of other imaging investigations to determine patient eligibility, added very shortly after the start of the trial, were prompted by feedback from a site visit to one centre and an early TSC meeting respectively.

Surgical form

It was anticipated that the surgical form (see *Appendix 21*) for the patient would be completed by a member of theatre staff. This form collected basic data on the type of surgery performed and key details relating to staffing and costs. Details on surgical techniques, such as surgical approach, were purposefully not collected. It was requested that the implant label was stuck on the back of the form. Also collected was information on unexpected procedures that occurred during the operation.

Inpatient episode form

The inpatient episode form (see *Appendix 22*) was completed at the conclusion of the patient's inpatient episode in the orthopaedic department. This form was important for collecting data on secondary outcomes such as mortality, surgical complications and treatment for shoulder-related complications and was the only source for collecting medical complications prior to discharge. It also provided a record of where the patient was discharged to and whether or not they had treatment for a serious newly diagnosed medical complication.

Physiotherapy treatment forms

The physiotherapy treatment log (see *Appendix 23*) was completed for each session of physiotherapy by the physiotherapist providing treatment to the patient. In designing this form, consideration was given to the normal record keeping of physiotherapists so that the completion of the form would require little additional effort. This included tick boxes for registering the use of basic categories of interventions, with space for recording other interventions, including whether or not there was any substantial difference from the ProFHER physiotherapy protocol. As indicated earlier, there was a question on whether or not the patient had performed their home exercises plus a text box for any comments. Key data on the physiotherapist's grade, the location of the session and session duration were collected for inclusion in the cost-effectiveness analysis. Piloting of draft data collection forms was undertaken at the lead centre by the lead physiotherapist (LG) to identify and resolve any problems before the start of the trial.

At the end of treatment, the physiotherapy treatment log – completion of treatment ('end of treatment') form (see *Appendix 24*) was usually completed by the physiotherapy contact at the centre in charge of co-ordinating data collection. This form captured the reasons for discharging a patient when treatment was and was not completed. The set of physiotherapy logs and the end of treatment form were then sent to the YTU.

One-year and 2-year follow-up forms

The designated person was responsible for the completion of the 1-year follow-up form (see *Appendix 25*) and the 2-year follow-up form (see *Appendix 26*) at the appropriate times for each patient. These forms were designed to collect information on complications, treatments and admissions that the patients may have had after their initial treatment and, for the 1-year form, their inpatient episode. Hospital admission for another fracture was recorded on both forms.

Adverse event reporting

Hospitals were provided with adverse event (reporting) forms (see *Appendix 27*) to record adverse events and to guide them in determining whether an event was a serious or a non-serious event. For a SAE, the form had to be completed and faxed to the trial manager within 24 hours of the local investigator becoming aware of the SAE. When the adverse event was not serious, the form had to be faxed within 5 days. Adverse events were judged also in terms of whether they were related or not to the patient

taking part in the trial (research procedures) and whether they were 'expected' or 'unexpected'. Unless the patient had died, the trial manager contacted the designated person 1 month after the original notification to request a completed review of adverse event form (see *Appendix 28*).

All adverse event and review forms were circulated to the chief investigator (AR) and two members of the TMG (HH and DT) for their feedback and consideration of whether or not any action should be taken, including reporting to the MREC. As per protocol, reports of all related and unexpected SAEs were submitted by the chief investigator to the MREC within 15 days of becoming aware of the event. The sponsor representative was also informed as these events arose. Details of all adverse events by treatment group were provided for review at DMEC meetings. The DMEC also reviewed adverse events after scrutiny by the TMG on a case-by-case basis. Count data of all adverse events, divided into non-serious and serious, were provided in the annual reports to the MREC. The funder (the National Institute for Health Research) was notified in the 6-monthly progress reports of the number of adverse events reportable to a regulatory body occurring in the last reporting period.

Trial exit form and withdrawal from follow-up

The trial exit form (see *Appendix 29*) was provided to facilitate notification and documentation when a trial participant did not want to continue with the trial or was no longer able to continue taking part in the trial. The key role of this form was to ensure that the trial participant or his or her family were not approached again. When appropriate, the trial manager checked with the participant whether or not he or she was happy for the collection of hospital follow-up data to continue.

Questionnaire on use of the ProFHER trial sling immobilisation leaflet and immobilisation

At the end of trial recruitment, a two-page questionnaire on sling use was sent to the designated person at each participating centre (see *Appendix 30*). The questions on the first page were aimed at gathering evidence on the implementation of a key standard of care by obtaining feedback on the use of the sling leaflet provided for the ProFHER trial or alternative leaflets on care during immobilisation provided to patients eligible for participation in the trial. The second page asked about immobilisation (type and duration) typically provided to non-surgical patients and to surgical patients after surgery who would have been eligible for the ProFHER trial. This was to obtain additional insights on this aspect of care, in particular on post-surgical immobilisation, for which we had not collected patient data.

Analysis of the questionnaire

The contents of the two pages of this questionnaire were processed and analysed separately. For the first page, draft analysis tables and criteria for assessment and interpretation of responses and alternative leaflets were generated by one TMG member (HH), who remained blinded to the completed questionnaires. These tables were completed by the trial manager (SB). Copies of any alternative leaflets were provided to HH, who summarised their contents.

For the second page, all of the responses were collated into a Microsoft Excel 2010 database (Microsoft Corporation, Redmond, WA, USA), which was then summarised by one person (HH). Feedback on the categorisation and reporting of the findings was obtained from the physiotherapy lead (LG).

Data checks and 'cleaning': specific actions

Data checks alerted us to recurrent problems with completion of two hospital forms during trial recruitment. Early on, missing data for the consent status form in terms of which treatment surgeons would advise for non-consenting patients was remedied by adding in a box to allow for uncertainty (see *Appendix 15*). The surgical form (see *Appendix 21*) proved troublesome in terms of data entry for

numbers of personnel involved, which were important data for the economic analysis. A follow-up visit to centres by the trial manager and advice in a monthly newsletter were strategies used to address this.

During data 'cleaning', two sources of problems were identified and corrective measures taken. A scrutiny of 'other reasons to exclude the patient' in the study eligibility forms revealed a very few cases in which lack of patient consent was recorded rather than non-eligibility. There were also some sequence and duplication problems in the recording of dates of physiotherapy sessions in the physiotherapy treatment logs (see *Appendix 23*).

Processing of data on complications and new therapy including surgery

A file of blinded data for each trial patient from the relevant fields in three forms [inpatient episode form (see *Appendix 22*), 1-year follow-up form (see *Appendix 25*) and 2-year follow-up form (see *Appendix 26*)] was compiled by the trial statistician (AK) to facilitate the independent screening and coding of complications and new therapy by two raters (HH and AR). Any disagreement on the processing, coding and categorisation of these items was resolved by discussion.

In the first stage, the complications and associated treatments were grouped into seven categories: complications related to the proximal humeral fracture and/or its treatment, gastrointestinal events, cardiac and peripheral vascular events, malignancy, falls and other fractures, respiratory events and 'other'. Processing of the first of these categories is described in the following section.

Surgical and other shoulder fracture-related complications, increased/new shoulder-related therapy and secondary surgery to the shoulder

A secondary file containing information on these complications and their treatment was generated to link each complication with the recorded treatment on the hospital form, when available, and to give information on the 'outcome' of each complication or treatment. This facilitated coding of the complications and the summarisation of treatment. It also facilitated the identification and coding of complications in which only treatment had been reported; each instance of this resulted in a new entry in the trial database by the statistician.

The categories used to code the complications were based primarily on the descriptions presented in the follow-up forms, with the decision taken to distinguish, when possible, between nonunion and malunion (*Table 3*). Additional categories were added, most notably 'post-traumatic stiffness', which included reported reference to 'frozen shoulder'. Rather than speculate, we coded all instances in which there was insufficient information as 'unclear'. Each complication was reported separately; thus, new entries in the definitive database were generated in the few cases in which two complications were described together.

Brief descriptions were generated for increased/new therapy for shoulder-related complications, types of shoulder surgery (based on categorisation by AR) and information on 'outcome'. Our interpretation was purposefully cautious. We did not complete missing entries for treatment but noted mention of treatment or attendance of follow-up clinics under 'outcome'.

We discarded entries for complications and treatment subsequent to definitive treatment for a complication. The information on these was instead summarised under 'outcome', which was categorised according to the definitions listed in *Table 4*.

TABLE 3 Descriptions of complications

Complications as described on forms	Complications as described in coding and presentation
Surgical site infection requiring treatment with antibiotics/further surgery	Surgical site infection
Haematoma formation at surgical site	Haematoma formation (not used)
Nerve injury	Nerve injury
Implant-related complication/failure	Implant-related complication/failure
Dislocation/instability	Dislocation/instability
Metalwork problems requiring further surgery	Metalwork problems
AVN of the humeral head	AVN
Nonunion/malunion requiring further treatment	Nonunion
	Malunion
Other	Post-traumatic stiffness
	Rotator cuff tear
	Complex regional pain syndrome
	Pain/severe pain
	Impingement
	Unclear

TABLE 4 Coding scheme for 'outcome' of a complication or treatment for a complication

'Outcome' category	Definition
No information	No information on outcome; however, other information on treatment and rehabilitation may be available. Although resolution of some complications would be very likely (e.g. metalwork impingement by metalwork removal; no reference to subsequent treatment for unresolved surgical site infection), explicit mention of this was required
Not treated	Clear indication that complication (usually malunion) was not treated nor is 'under review'
Resolved	Explicit statement that complication resolved. This includes cases in which physiotherapy was required to resolve initial problems after secondary surgery
Resolving	Explicit statement that complication was resolving (final resolution was not assumed)
Under review	Clear indication that original (primary) complication was being monitored/investigated, with or without reference to future new treatment/surgery
Unresolved	Consequence of definitive treatment (mainly surgery) in which either the original problem persisted or a new problem arose
Unresolved – further surgery	As above, but further surgery reported

Medical complications prior to discharge

Data for this secondary outcome were collected from the inpatient episode form only (see *Appendix 22*). These were grouped in the categories described on this form, with any appearing in the two 'other' categories ('other hospital-acquired infection' and 'other serious event') being placed into one of six groups: gastrointestinal events, cardiac and peripheral vascular events, malignancy, falls and other fractures, respiratory events and others.

Treatment for serious newly diagnosed medical complications

Data on these complications were collected only when they were first recorded, with any subsequent reports of recurrence discarded. Each newly diagnosed complication was placed into one of the six aforementioned groups.

Further fractures

Data on subsequent fractures were collected on a most appropriate mention basis, with other reports relating to the same fracture discarded. A secondary table with the descriptions of the fractures and source key information was compiled and the fractures were independently grouped by the two raters into the same four categories used for grouping fractures reported as adverse events (see following section): shoulder, wrist, hip and other.

Adjusting the database for complications and treatment

After screening and coding, the database was amended to reflect the decisions made and coding of the various complications and treatments. One rater (HH) coded the discarded entries, which for reporting purposes were honed down to four categories by AK: repeat information, subsequent/additional information on treatment for a complication, irrelevant content and, for further fracture data, routine follow-up.

Adverse events classification

A protocol was developed to inform the reporting and classification of adverse events. This established the following:

- Appraisal of adverse events would be based only on the adverse event reporting forms and thus would be independent of information on complications from hospital forms such as the 1-year follow-up form.
- Events reported after 2 years' follow-up would not be included.
- We would present separate results, divided by treatment group, for SAEs and non-SAEs. In these two categories, we would report the number of participants who had an adverse event and, as some participants had multiple events, also the number of adverse events.
- We would report the classification of the adverse event based on the judgement of the principal investigator at the centre reporting the adverse event of whether or not the event was (a) related and (b) expected.
- Two reviewers (HH and AR) would independently group, without knowledge of treatment allocation, the adverse events based on the categories used on the hospital forms. Facilitated by the trial manager (SB), consensus would be reached through discussion when there was disagreement on the categories and on the brief descriptions of the events added to the data file by the reviewers. *Box 1* presents the agreed categories.

BOX 1 Categories for coding adverse events

1. Shoulder-related complications (e.g. infection, implant complication, frozen shoulder),^a which are subgrouped according to whether or not there is:
 - i. surgery to the shoulder (if there is any mention of surgery to the shoulder)
 - ii. non-surgical treatment/investigation of the shoulder.
2. Medical complications/conditions that do not involve the shoulder, which are subgrouped according to whether or not it is:
 - i. non-serious
 - ii. serious: non-fatal
 - iii. serious: fatal.
3. Any other fractures, which are subgrouped according to whether or not they involve the:
 - i. shoulder
 - ii. wrist
 - iii. hip
 - iv. other.

a This should exclude a second shoulder fracture as this is not a complication that is related to the original injury.

Baseline radiographs: processing and quality assessment

Initial processing of the radiographic images, all of which were in digital format, involved checks on blinding (removal of patient and hospital details), labelling with the correct patient identifier and ensuring that the images could be accessed and transferred to a YTU computer and CD. The designated person and/or the contact in the radiography department was contacted to resolve difficulties at individual centres. For blinding purposes, before sending out for quality checks (see following paragraph) and eventual classification, the sets of radiographs were renumbered using a three-digit code (e.g. 039), with letters used to label each of the radiographs within a set (e.g. 039a, 039b).

At various times during trial recruitment, and subsequently to complete the process, there was a review of the quality of the copies of the radiographic images for each trial participant provided by trial centres. This aimed to facilitate the Neer classification of the study fractures by two independent assessors by ensuring that such images were available and of sufficient quality for this purpose. We undertook a two-phased assessment for each centre whereby the quality of the radiographs for the first five participants at each centre was independently assessed by two orthopaedic surgeons, one of whom was the chief investigator (AR) and other of whom was a principal investigator (JC); the assessment of the radiographs of subsequent participants from each centre was carried out by AR only. When there were concerns about the quality of the radiographs being provided, this was to be raised by AR with the principal investigator of the participating centre to see whether or not better copies could be obtained. The criteria that were used to judge the quality of the radiographs were:

- Are there at least two projections in planes perpendicular to each other? (yes/no)
- Are the proximal humerus and glenohumeral joint seen on each projection? (yes/no)
- Is it possible on the two views to clearly identify the shaft, greater tuberosity, lesser tuberosity, head of the humerus and the glenohumeral joint? (yes/no).

The review also included monitoring for a clear breach of the main inclusion criterion; thus, fractures should involve the surgical neck. The trial manual alerted centres to this intention to 'monitor the eligibility of patients entering the trial', stating also that any findings from the interim review would 'not be referred back to individuals but applied generally'. However, there was the potential to report to the TSC for consideration of discontinuation of recruitment at a site where there were ongoing problems with poor-quality images and serious concerns about patient eligibility.

On discussion with the TSC of the interim findings on radiograph quality, the independent orthopaedic surgeon member of the TSC (Professor Andrew Carr) volunteered to act as a third rater for resolving disagreements in the final stage of the quality audit. This staged audit identified 46 radiograph sets in which at least one out of the three quality criteria had not been met, as judged either by AR alone ($n = 28$) or, after arbitration by AC, by two independent raters ($n = 18$). This meant that 18% of the baseline radiograph sets were likely to present future difficulties for the independent Neer classification of fractures.

Post-recruitment survey of hospital radiographers

On consideration of the interim findings on radiograph quality, which showed that, at that stage, over one-quarter of the radiographic sets did not meet our quality criteria, the TSC advised that we contact the participating centres to obtain feedback on this aspect of the trial. We did this by sending out a survey in December 2011 to the named radiographers of the individual sites (see *Appendix 31*). As stated in our covering letter, this survey aimed to obtain their specialist feedback on a key part of the care pathway for people with proximal humeral fractures. Thus, we had extended the scope of the questionnaire to obtain a better understanding of the issues relating to imaging for these patients, including the routine use of computerised tomography (CT) scans. The results of this survey are reported in *Appendix 32*.

In connection with the issues relating to assessing radiograph quality for suitability for the Neer classification, the survey feedback prompted two actions. First, the wording of the second two criteria was revised to make clearer what was required; these were piloted subsequently, with the final versions appearing in the proforma used for the Neer classification. Second, based on suggestions for other quality criteria, another question was included in the proforma that asked about overall image quality, taking into consideration exposure and patient positioning.

Preparations for the independent assessment and classification of radiographs based on the Neer classification

Key to the Neer classification of the radiographs is the facility to assess whether or not there is ≥ 1 cm displacement of parts, with clear delineation of the edge of the bones. (The assessment of angular displacement is fortunately not size or scale independent.) Hence, to measure this displacement, a life-sized image or a scale is required. In retrospect it was discovered that, for some Picture Archiving and Communications Systems (PACS), the scale as well as patient details were removed when the images were anonymised. On investigating this, we discovered that it could be remedied by saving the radiograph image as a picture and then removing the patient details manually. However, this may not be an acceptable procedure at some hospitals and is still subject to scaling errors. We learnt, through discussions with a radiographer (Brian Cox) at James Cook University Hospital, Middlesbrough, that a calibration object such as a scaling ball, placed in an appropriate location, is required at image acquisition to enable the accurate measurement of scale. We identified two albeit suboptimal approaches of limited applicability to assist the assessment of linear displacement in our study sample:

- For a limited number of radiographs for which standard left/right markers are used, the stems of the 'L' and 'R' are 9 mm. It was therefore possible to measure the stem of the 'L' or 'R' marker on the image to adjust for the scale when measuring for displacement.

- Software such as Microsoft Office Picture Manager (Microsoft Corporation, Redmond, WA, USA) could be used to allow a JPEG file to be opened and viewed at the original size that it was saved at. (This still may not be a true result if the image was manipulated in some way before it was saved.)

In addition to being able to present the best-quality images available with the right scale to accurately assess displacement, we also considered what training was necessary for the two raters who would be reviewing the radiographs. We therefore revisited a systematic review²² and various individual reports of studies investigating inter-rater and intra-rater reliability for classifying proximal humeral fractures using the Neer system, noting in particular the approaches taken for assessing displacement and training.^{19,63–65} We also approached Stig Brorson (Herlev University Hospital, Herlev, Denmark) for details of the training programme used in his RCT, which showed that interobserver agreement is improved with training.⁶⁵ Brorson reported that the authors of the RCT had chosen to include only material from Neer's original articles (22 and 29 March 2012, personal communication).^{18,66} Eighteen illustrations from the first article and two from the second article were reproduced and incorporated into two 45-minute teaching sessions. We decided against adopting this approach as we considered that using actual ProFHER trial images was more relevant for facilitating the training of the two raters. This decision was consistent with our aim to have continuity in the images available for the decision-making process – from those seen by the recruiting surgeon to those seen by the independent assessor.

When considering which of the baseline radiographic images should be available for the two raters to review for the Neer classification we decided that it would be in excess to have more than four images per set of radiographs. This is because we expected that some images were poorer-quality duplicates of standard views or uninformative radiographs that would not contribute to the Neer classification. On provision of the baseline radiographs and a table listing all of the radiograph image sets, AR selected the best-quality projection (based on prespecified criteria) in each perpendicular plane. The duplicates in these planes that were of poorer quality were removed, with the reasons for exclusion being recorded. The decisions that AR made were checked by two others (SB and HH), who sought clarification when necessary.

We developed a protocol, training presentation and data collection tool (proforma) to aid in the description and assessment of the quality of each set of radiographs and the Neer classification of the study fractures. Throughout, the purpose of quality assessment was to establish if the sets of radiographs were adequate to permit their classification and not if they were adequate for use in clinical practice. As well as collecting data on the displacement of structures according to the Neer classification, the proforma facilitated the recording of 'involvement' (any indication of a fracture), a fully displaced surgical neck fracture (bone parts/fragments do not overlap), whether or not the head segment was in varus or valgus and whether or not the fracture was impacted.

The training presentation and data collection process were piloted with the help of two senior orthopaedic registrars at the James Cook University Hospital, Middlesbrough, on 6 August 2012. For this pilot, AR selected 10 sets of patient radiographs based on the following criteria: with and without adequate markers (i.e. with markers that were added to the cassette at the time of image acquisition and with markers that were added using PACS); a selection of two, three and four views (which for the last should include a couple with similar anteroposterior views); and a choice of two- compared with three- and four-part fractures so that some would have tuberosity involvement. After AR had delivered the presentation, which described the objectives of the Neer classification study and provided training in the interpretation and classification of radiographs, he and the two registrars independently classified 10 sets of radiographs according to the Neer classification using the proforma. A ruler and goniometer were made available to the raters, who were advised to use these. The trial manager (SB) acted as a general facilitator, including timing the various processes and documenting the proceedings, including the inter-rater differences after each of the two batches of five assessments. This enabled the three raters to discuss their differences and decide on the Neer classification for each fracture. The insights from the pilot led to some adjustment of the proforma (see *Appendix 33* for the final version), clarification of the advice on interpretation,

development of a briefing document on the Neer classification of radiographs and a form to record discrepancies between the two raters and their resolution and development of a realistic timetable for the main assessment of the baseline radiographs and classification of the study fractures according to Neer.

The main Neer classification of the radiographs was undertaken by two independent consultant orthopaedic surgeons (from the UK) who were experienced in treating proximal humeral fractures and who had comparable experience to that of surgeons in the ProFHER trial. This differed from our original intention to use an independent panel of musculoskeletal radiologists or orthopaedic surgeons with specific expertise in the Neer classification. However, it was in keeping with the pragmatic design of the trial and, from our experience of the audit described earlier and the literature, it was clear that involving more raters does not improve agreement. When agreeing to perform this role, both surgeons confirmed that they had the written support of their clinical director to commit to the requirements for participation. This comprised attendance at the training day (held on 17 January 2013), protected clinic time (five sessions: 20 hours) to undertake assessment of the radiographs and up to a day to achieve consensus when required. They also confirmed that they would 'complete the assessment of radiographs for the whole study population according to the agreed procedures and undertake to deliver a definitive categorisation of the fractures, if necessary via an adequately-documented consensus process, within the timeline for this study' (letter from AR to the two orthopaedic surgeons setting out the terms, 6 November 2012).

The same format and number of sets of radiographs were used for the main study training day as in the pilot. Thus, after the presentation by AR with examples, the two independent surgeons looked at the first five sets of radiographs and reviewed discrepancies; this was followed by assessment of the second five sets of radiographs, with a review of discrepancies. Subsequently, the surgeons returned copies of their completed data collection forms to SB, who collated the results and returned a table indicating where there were differences in the Neer classification for individual fractures. The two surgeons then met up to resolve these differences and document the decisions behind each of the final verdicts.

Chapter 4 Results: baseline characteristics of patients, delivery of interventions and return of questionnaires and other data

Patient recruitment

The first patient was recruited on 17 September 2008, which was during the short 3-month pilot phase, and the 250th patient was recruited on 13 April 2011. Thus, patients were recruited over a 31-month period; the trial stopped when the recruitment target was met. *Figure 4* shows the monthly recruitment and accrual set in the context of the revised projections submitted with our successful application for an extension to recruitment from the original 18 months (from the official start time of the main trial recruitment of 1 October 2008) to 36 months.

Ultimately, 32 out of 35 UK-based centres that had been set up to recruit after approvals had been received recruited a patient. One of the three centres that did not recruit had started screening but identified only ineligible patients. A list of the participating centres, and the primary care trusts for which R&D approval for follow-up was obtained, is given in *Appendix 34*. A full list of health-care collaborators at the participating centres is presented in *Appendix 35*.

The 35 participating centres recruited a median of five patients [interquartile range (IQR) 2–9 patients]. The lead centre, James Cook University Hospital, Middlesbrough, recruited 43 patients and achieved the original centre recruitment target of one patient per month. *Figure 5* shows the distribution of number of patients recruited by centre, ordered by when each site was set up to recruit.

Location in hospital of patients at recruitment

The location (fracture clinic or ward) at which each patient was recruited from was recorded at the time of randomisation. These data showed that 36 (14%) patients had been admitted and the rest were outpatients attending a fracture clinic. Of these 36, 15 were randomised to 'surgery' (12% of the 125 patients randomised to surgery) and 21 were randomised to 'not surgery' (17% of the 125 patients randomised to 'not surgery').

Population sample

Based on the return of correctly completed study eligibility forms and consent status forms, 1250 patients with a suitable fracture were identified, of whom 563 (45%) were considered eligible; of these 563 patients, 250 (44%) consented to take part in the study. Thus, of 1250 patients identified with a suitable fracture, 250, or 20%, were enrolled into the study. This figure compared favourably with our original expectation to recruit only 11% of patients with fracture types suitable for inclusion in the trial.

Reasons for exclusion

The reasons for excluding patients (790 reasons for 687 ineligible patients), as listed on the study eligibility form (see *Appendix 5*), are shown in *Table 5*.

The numbers of ineligible patients and the reasons for exclusion at individual centres and overall during the recruitment period were monitored on a regular basis by the TMG. In particular, there was an ongoing review of 'other' reasons for exclusion (option j), which might indicate potential problems at centres, including lack of surgeon equipoise ('protocol violation'). These were followed up when deemed necessary.

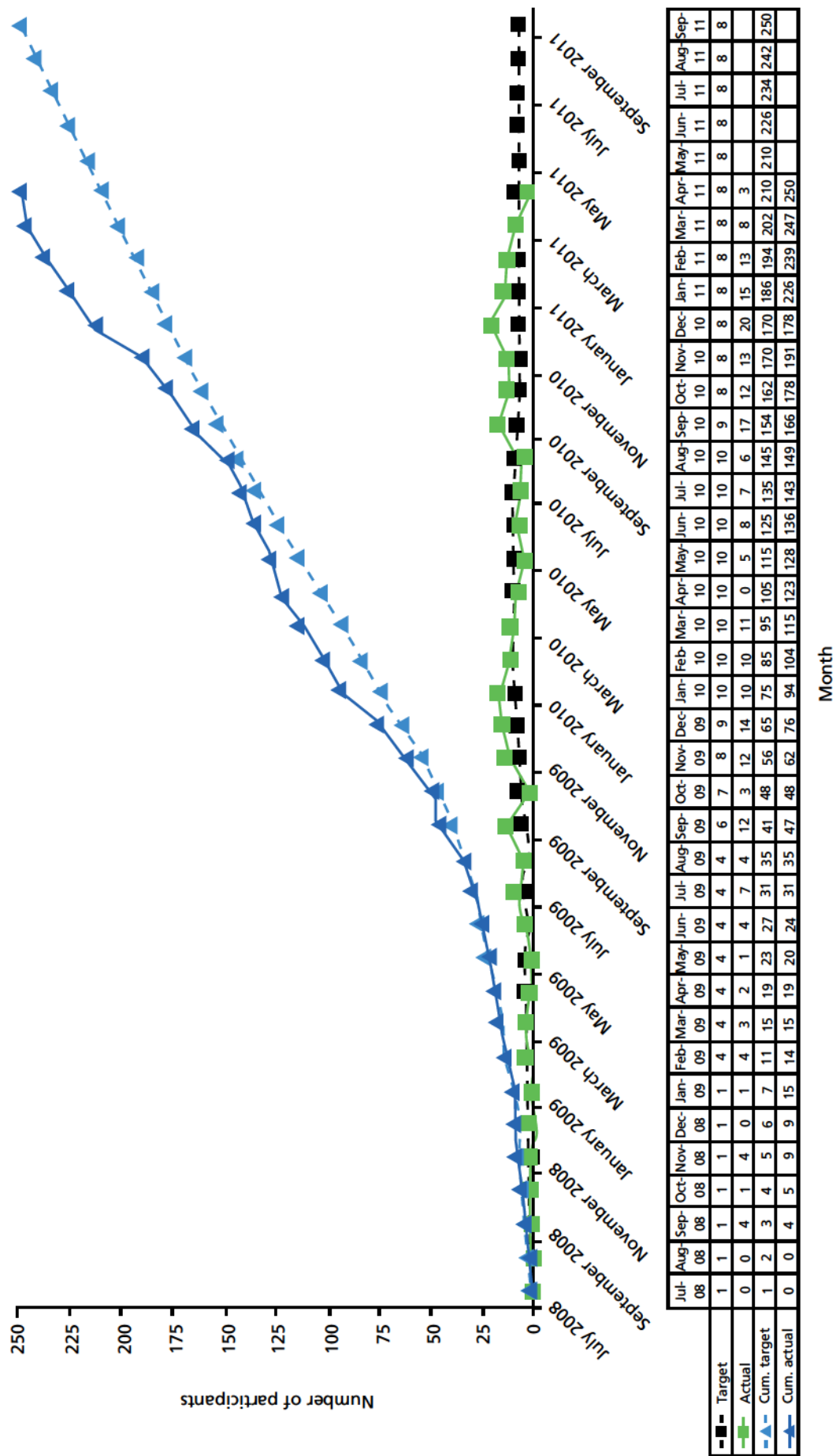


FIGURE 4 Patient recruitment and accrual in the ProFHER trial. Cum., cumulative.

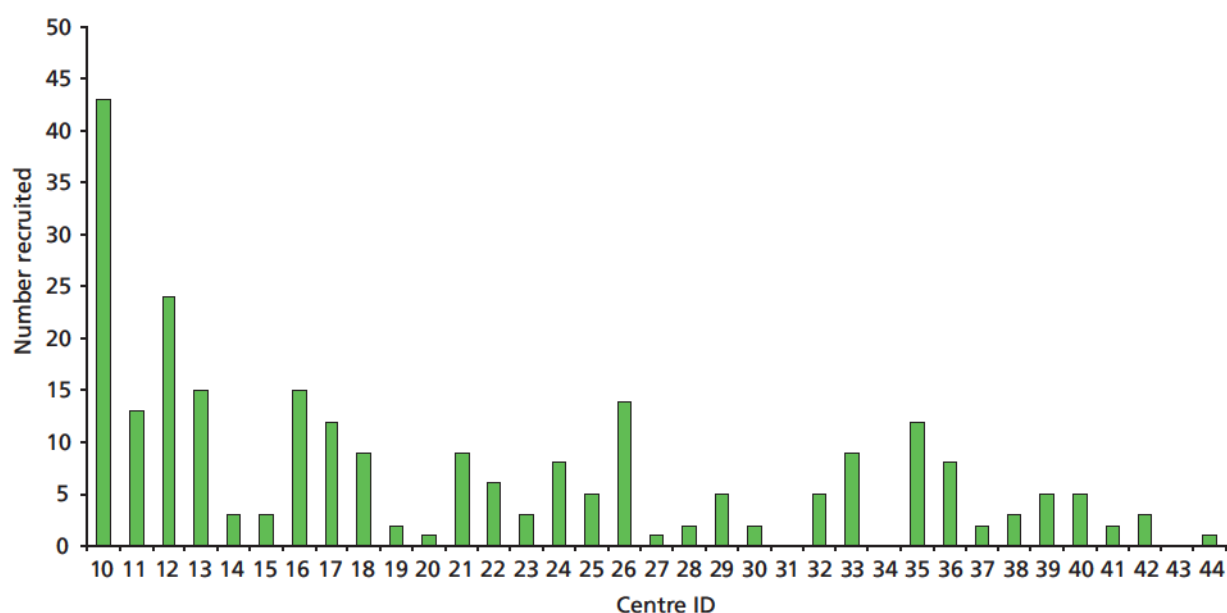


FIGURE 5 Patient recruitment by centre listed in order of when the centre was set up to recruit.

TABLE 5 Reasons listed for exclusion

Reason for exclusion (more than one reason may apply)	Number excluded
Associated dislocation of the injured shoulder joint	101
Open fracture	2
Other upper limb fractures	72
Clear indication of surgery	87
Comorbidities that preclude surgery or anaesthetic	179
Pathological fracture	5
Terminal illness	5
Patient not resident in hospital catchment area	28
Lack of mental capacity	116
'Other' reason	195 ^a

a For 159 of these no other reason was given.

Exclusions because of a patient's lack of mental capacity to understand the trial and comply with trial procedures (option i) were selected for special focus by the TSC, which was provided with regular reports of the percentage of overall exclusions for this item. There was a consistent rate of between 14% and 18% of the total reasons given at each monitoring period and the initial concern regarding inappropriate exclusion based on an excess of people excluded for mental incapacity was allayed.

At the end of recruitment and collection of all trial eligibility forms, two raters (AR and HH), who were blind to centre, independently assessed the 'other' reasons listed for excluding patients (option j; see *Appendix 5*). They concluded that 117 patients (17% of a total of 687 ineligible patients) were excluded because of a lack of equipoise. As these 117 patients were otherwise eligible for the trial, a more appropriate statistic is that 17% of 680 (563 + 117) patients who were actually 'eligible' for the trial were not approached for consent because of a lack of surgeon equipoise. This is examined in more detail in *Chapter 6*.

Patient consent

The percentage of eligible patients who consented to take part in the trial varied substantially in the 32 centres that recruited a patient, from 20% to 100% (*Figure 6*). Patient treatment preference data were collected but otherwise, patients were purposefully not asked for their reasons for not consenting. It is likely that there are various reasons for this variation in the rate of consent. It should be noted that all of the designated persons involved in consenting were provided with advice and training on the process involved but still had variable skills and experience in recruiting patients into a trial.

Participant flow

The flow of patients from screening to the 2-year follow-up as well as the data used for the primary analysis are summarised in a Consolidated Standards of Reporting Trials (CONSORT) flow diagram in *Figure 7*. Of the 250 patients randomised in equal numbers to surgery or not surgery, 215 (86%) were available for analysis at 2 years' follow-up (85% in the surgical arm, 87% in the non-surgical arm). The trial was designed to allow for a 20% dropout rate.

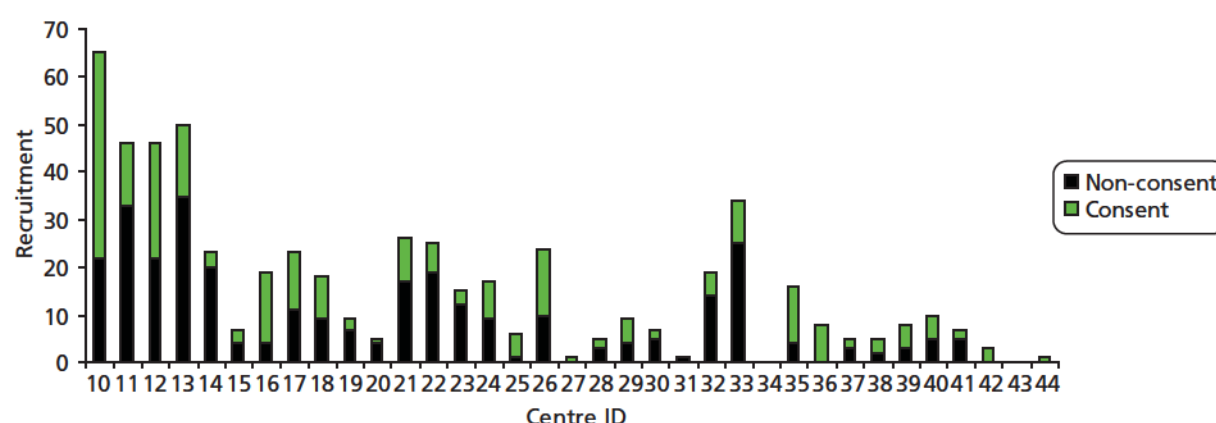


FIGURE 6 Non-consenting and consenting patients by centre.

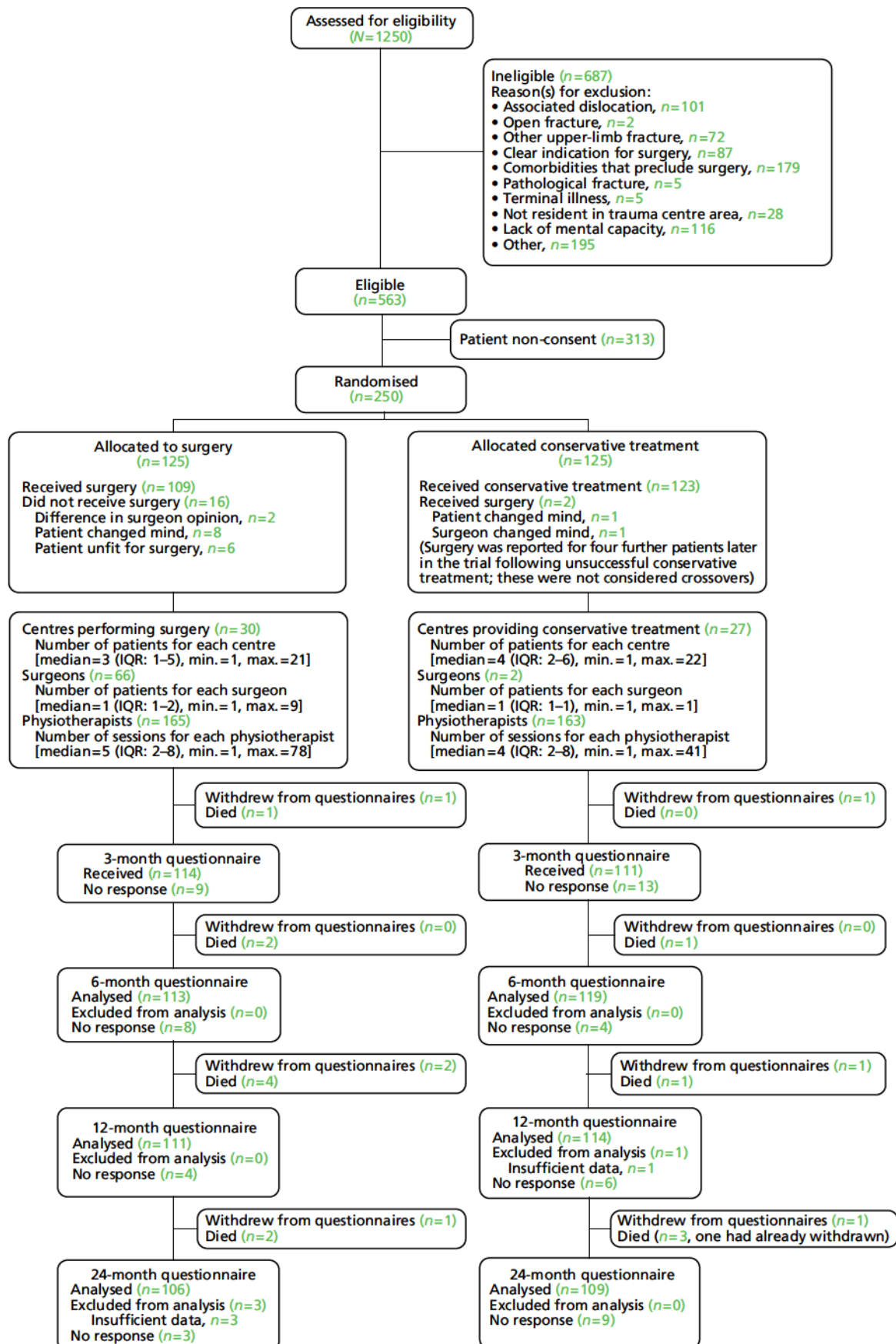


FIGURE 7 Participant flow through the trial. Max., maximum; min., minimum.

Baseline characteristics: all screened patients

The characteristics of the screened, ineligible, eligible consenting and eligible non-consenting patients at baseline are presented in *Table 6*. Eligible patients tended to be younger, presenting slightly later following injury, and were more likely to have fractures that involved tuberosities than ineligible patients. There were no marked differences between consenting and non-consenting patients.

Baseline characteristics of randomised patients

Consenting patients were stratified by tuberosity involvement (yes or no) at randomisation. Their baseline demographics and fracture details by treatment group are given in *Tables 7–9*. The treatment groups appeared to be balanced for these characteristics, including Neer category, with the exception of smoking status. Patients in the non-surgical group were more likely to be smokers (smoking status was included in a sensitivity analysis for the primary outcome; see *Chapter 5*). The same observations applied to the baseline characteristics of the 215 patients with OSS data at 24 months (see *Tables 7 and 8*).

TABLE 6 Baseline characteristics (different populations)

Characteristic	Screened (<i>n</i> = 1250)	Ineligible (<i>n</i> = 687)	Eligible (<i>n</i> = 563)	
			Non-consenting (<i>n</i> = 313)	Consenting (<i>n</i> = 250)
Gender, <i>n</i> (%)				
Male	318 (25.4)	185 (26.9)	75 (24.0)	58 (23.2)
Female	924 (73.9)	496 (72.2)	236 (75.4)	192 (76.8)
Missing	8 (0.6)	6 (0.9)	2 (0.6)	–
Age (years)				
<i>n</i>	1192	644	298	250
Mean (SD)	69.03 (14.85)	70.45 (16.41)	68.48 (13.04)	66.01 (11.94)
Median (min., max.)	71.12 (16.09, 106.19)	73.61 (16.09, 106.19)	70.56 (18.42, 98.39)	66.88 (24.63, 92.04)
Time since injury (days) ^a				
<i>n</i>	1250	687	313	250
Mean (SD)	5.21 (4.92)	5.03 (5.01)	5.19 (4.72)	5.74 (4.89)
Median (min., max.)	3 (0, 21)	3 (0, 21)	4 (0, 21)	4 (0, 21)
Tuberosity involvement, <i>n</i> (%)				
Tuberosity not involved (or missing data)	417 (33.4)	276 (40.2)	84 (26.8)	57 (22.8)
Greater tuberosity	536 (42.9)	255 (37.1)	162 (51.8)	119 (47.6)
Lesser tuberosity	62 (5.0)	36 (5.2)	16 (5.1)	10 (4.0)
Greater and lesser tuberosity	235 (18.8)	120 (17.5)	51 (16.3)	64 (25.6)

Max., maximum; min., minimum.

^a This is consistent with the inclusion criterion for presentation of patients within 3 weeks of their injury.

TABLE 7 Baseline characteristics (demographics) at randomisation and 24 months' follow-up

Characteristic	All randomised		Those with primary outcome (OSS) data at 24 months	
	Surgery (n = 125)	Not surgery (n = 125)	Surgery (n = 106)	Not surgery (n = 109)
Gender, n (%)				
Male	28 (22.4)	30 (24.0)	26 (24.5)	25 (22.9)
Female	97 (77.6)	95 (76.0)	80 (75.5)	84 (77.1)
Age (years)				
n	125	125	106	109
Mean (SD)	66.60 (11.80)	65.43 (12.09)	66.18 (11.1)	65.79 (11.97)
Median (min., max.)	67.42 (27.04, 92.04)	66.12 (24.63, 89.02)	66.67 (37.09, 87.76)	66.77 (31.33, 89.02)
Age (group), n (%)				
< 65 years	51 (40.8)	57 (45.6)	45 (42.5)	50 (45.9)
≥ 65 years, n (%)	74 (59.2)	68 (54.4)	61 (57.5)	59 (54.1)
Ethnicity, n (%)				
White	124 (99.2)	125 (100.0)	105 (99.1)	109 (100.0)
Black	1 (0.8)	0 (0.0)	1 (0.9)	0 (0.0)
Asian	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Chinese	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Other	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Education, n (%)				
No formal qualifications	66 (52.8)	68 (54.4)	53 (50.0)	57 (52.3)
Some qualifications/no degree	47 (37.6)	43 (34.4)	43 (40.6)	38 (34.9)
Degree or higher	12 (9.6)	14 (11.2)	10 (9.4)	14 (12.8)
Employment, n (%)				
Part-time	12 (9.6)	7 (5.6)	12 (11.3)	7 (6.4)
Full-time	17 (13.6)	22 (17.6)	16 (15.1)	19 (17.4)
Self-employed	1 (0.8)	3 (2.4)	1 (0.9)	3 (2.8)
Retired	78 (62.4)	82 (65.6)	64 (60.4)	72 (66.1)
Not employed but seeking work	3 (2.4)	1 (0.8)	3 (2.8)	1 (0.9)
Other	12 (9.6)	9 (7.2)	9 (8.5)	6 (5.5)
Missing	2 (1.6)	1 (0.8)	1 (0.9)	1 (0.9)
Diabetes, n (%)				
Yes	18 (14.4)	13 (10.4)	15 (14.2)	11 (10.1)
No	106 (84.8)	111 (88.8)	90 (84.9)	97 (89.0)
Missing	1 (0.8)	1 (0.8)	1 (0.9)	1 (0.9)

continued

TABLE 7 Baseline characteristics (demographics) at randomisation and 24 months' follow-up (*continued*)

Characteristic	All randomised		Those with primary outcome (OSS) data at 24 months	
	Surgery (n = 125)	Not surgery (n = 125)	Surgery (n = 106)	Not surgery (n = 109)
Smoking status, n (%)				
Yes	24 (19.2)	40 (32.0)	20 (18.9)	33 (30.3)
No	96 (76.8)	81 (64.8)	82 (77.4)	72 (66.1)
Missing	5 (4.0)	4 (3.2)	4 (3.8)	4 (3.7)
Steroid use, n (%)				
Yes	6 (4.8)	7 (5.6)	6 (5.7)	6 (5.5)
No	118 (94.4)	116 (92.8)	100 (94.3)	102 (93.6)
Missing	1 (0.8)	2 (1.6)	0 (0.0)	1 (0.9)
Health Status (EQ-5D index)				
n	123	121	104	106
Mean (SD)	0.43 (0.37)	0.38 (0.37)	0.43 (0.35)	0.35 (0.36)
Median (min., max.)	0.59 (−0.36, 1) ^a	0.26 (−0.35, 1) ^a	0.59 (−0.35, 1) ^a	0.26 (−0.35, 1) ^a
Max., maximum; min., minimum.				
a An EQ-5D score of 1 (perfect health) was reported for 11 surgery group patients and 13 non-surgery group patients. It is possible that this applied to their pre-fracture status.				

TABLE 8 Baseline characteristics (fracture data) at randomisation and 24 months' follow-up

Characteristic	All randomised		Those with primary outcome (OSS) data at 24 months	
	Surgery (n = 125)	Not surgery (n = 125)	Surgery (n = 106)	Not surgery (n = 109)
Time since injury (days)				
n	125	125	106	109
Mean (SD)	5.78 (4.90)	5.69 (4.89)	5.81 (5.00)	5.69 (4.82)
Median (min., max.)	4.00 (0.00, 19.00)	4.00 (0.00, 21.00)	4.00 (0.00, 19.00)	4.00 (0.00, 21.00)
Affected shoulder, n (%)				
Left	57 (45.6)	68 (54.4)	46 (43.4)	58 (53.2)
Right	68 (54.4)	57 (45.6)	60 (56.6)	51 (46.8)
Tuberosity involvement, n (%)				
Yes	99 (79.2)	94 (75.2)	85 (80.2)	83 (76.1)
No	26 (20.8)	31 (24.8)	21 (19.8)	26 (23.9)
Tuberosity involvement, n (%)				
Tuberosity not involved or missing	26 (20.8)	31 (24.8)	21 (19.8)	26 (23.9)
Greater tuberosity	58 (46.4)	61 (48.8)	51 (48.1)	56 (51.4)
Lesser tuberosity	7 (5.6)	3 (2.4)	5 (4.7)	2 (1.8)
Greater and lesser tuberosity	34 (27.2)	30 (24.0)	29 (27.4)	25 (22.9)

TABLE 8 Baseline characteristics (fracture data) at randomisation and 24 months' follow-up (*continued*)

Characteristic	All randomised		Those with primary outcome (OSS) data at 24 months	
	Surgery (n = 125)	Not surgery (n = 125)	Surgery (n = 106)	Not surgery (n = 109)
Radiographs used, n (%)				
Anteroposterior view only	5 (4.0)	17 (13.6)	3 (2.8)	15 (13.8)
Axillary view only	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Scapular Y-lateral view only	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Anteroposterior and axillary view	42 (33.6)	43 (34.4)	33 (31.1)	39 (35.8)
Anteroposterior view and scapular Y-lateral view	35 (28.0)	26 (20.8)	31 (29.2)	22 (20.2)
Axillary view and scapular Y-lateral view	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Anteroposterior and axillary view and scapular Y-lateral view	40 (32.0)	36 (28.8)	36 (34.0)	30 (27.5)
Missing	3 (2.4)	3 (2.4)	3 (2.8)	3 (2.8)
Previous fractures, n (%)				
Yes	33 (26.4)	33 (26.4)	27 (25.5)	30 (27.5)
No	92 (73.6)	90 (72.0)	79 (74.5)	77 (70.6)
Missing	0 (0.0)	2 (1.6)	0 (0.0)	2 (1.8)
Previous surgery for fractures, n (%)				
Yes	8 (6.4)	12 (9.6)	6 (5.7)	10 (9.2)
No	23 (18.4)	21 (16.8)	19 (17.9)	20 (18.3)
Missing	2 (1.6)	0 (0.0)	2 (1.9)	0 (0.0)
No previous fractures	92 (73.6)	92 (73.6)	79 (74.5)	79 (72.5)
Shoulder dominance, n (%)				
Yes	67 (53.6)	61 (48.8)	57 (53.8)	55 (50.5)
No	56 (44.8)	62 (49.6)	48 (45.3)	52 (47.7)
Missing	2 (1.6)	2 (1.6)	1 (0.9)	2 (1.8)
Injury mechanism, n (%)				
Fall/trip from standing height or less	90 (72.0)	96 (76.8)	77 (72.6)	84 (77.1)
Fall downstairs/steps or from a height	18 (14.4)	17 (13.6)	15 (14.2)	15 (13.8)
Other	15 (12.0)	9 (7.2)	12 (11.3)	7 (6.4)
Missing	2 (1.6)	3 (2.4)	2 (1.9)	3 (2.8)

Max., maximum; min., minimum.

TABLE 9 Baseline characteristics: types of fractures according to the Neer classification

Final Neer classification categories (see <i>Appendix 2</i>) ^a		Allocation, n (%)	
		Surgery (n = 125)	Not surgery (n = 125)
1 ^b	Neer one part: undisplaced ^c surgical neck	9 (7.2)	9 (7.2)
3	Neer two part: surgical neck	60 (48.0)	59 (47.2)
4 ^b	Neer two part: greater tuberosity	5 (4.0)	3 (2.4)
5 ^b	Neer two part: lesser tuberosity	0 (0.0)	1 (0.8)
8	Neer three part: surgical neck + greater tuberosity	45 (36.0)	45 (36.0)
9	Neer three part: surgical neck + lesser tuberosity	1 (0.8)	0 (0.0)
10 ^b	Neer three part: anterior dislocation + greater tuberosity	0 (0.0)	2 (1.6)
12	Neer four part: surgical neck + greater tuberosity + lesser tuberosity	5 (4.0)	6 (4.8)

a The categories in bold (categories 3, 8, 9 and 12) are those that fully meet the ProFHER trial inclusion criteria.
b These categories are outside the expected categories for the trial.
c Although these fractures were rated as 'undisplaced' for the purposes of the Neer classification, they were all displaced fractures in which the degree of displacement was insufficient for categorisation as a 'part'.

Table 9 shows the results of the independent classification of the fractures according to Neer (see *Chapter 3*). The categories in bold (categories 3, 8, 9 and 12) are those that fully meet the ProFHER inclusion criteria, accounting for 221 of the 250 participants (88%). Consistent with the pragmatic study design, there were other fractures involving the surgical neck that did not meet the Neer criteria but which were sufficiently displaced for inclusion of these patients into the trial because of their potential suitability for surgery. The two surgeons undertaking the classification confirmed that all 250 fractures met the main study criterion, that is, that the fracture involved the surgical neck. Only category 10, involving humeral head dislocation, is a protocol violation. It was noteworthy that both fracture dislocations were categorised at the consensus meeting; before this, the two independent surgeons had assigned each of these fractures to category 8. This indicates some uncertainty and difficulty in the classification of these two fractures.

For the purposes of the sensitivity analysis, secondary to a subgroup analysis of tuberosity involvement, categories 1 and 3–5 were combined to form the one- or two-part fractures group, with the other four categories forming the three- or four-part fractures group (see *Chapter 5*). More detailed results and analysis of the baseline radiographs and the Neer classification results are presented in *Chapter 6*.

Interventions: surgery

Of the 125 patients allocated to surgery, 109 received surgery and 16 (13%) were treated non-surgically for the following reasons: difference in opinion of surgeon ($n = 2$), patient changed mind ($n = 8$) and patient unfit for surgery ($n = 6$).

Surgery took place on average 4.8 days (median 4 days; range 0–21 days) from randomisation and an average of 10.4 days (median 9 days; range 1–33 days) from the date of injury.

Operating surgeon

It was expected that, as part of good standard care, surgery for these types of fractures would usually be carried out by a consultant who was experienced in operating on such fractures. The 109 operations in the surgery group were carried out by 66 different surgeons (operations: median 1, IQR 1–2, range 1–9). The grade of the operating surgeons is summarised in *Table 10*.

TABLE 10 Staff grade of operating surgeons ($n = 109$ participants)

Operating surgeon	n (%)	Notes
Consultant	89 (82)	
Consultant or registrar	8 (7)	Both were present in the theatre but we were unable to determine the status of the named surgeon
Registrar (consultant present in theatre)	5 (5)	Both were present in the theatre – the registrar was the operating surgeon
Registrar or fellow	7 (6)	All were confirmed to be senior registrars/specialist fellows. The availability of the consultant was confirmed in all cases

The standard short curricula vitae (CVs) that surgeons provided for the purposes of obtaining site R&D approval were scrutinised to ascertain specialty (how the surgeon described him- or herself) and appointment history (appointments within the previous 5 years). These CVs provided very little insight on specialty (most stated 'orthopaedics' or 'orthopaedics and trauma'). Of the 89 operations carried out by consultants, 46 were conducted by surgeons who had been consultants for at least 5 years, 18 were conducted by surgeons who had become consultants within the past 5 years and 25 were conducted by consultant surgeons for whom CVs were not available.

As expected in our protocol, the majority ($n = 89$; 82%) of operations were performed by consultant surgeons, with a further 13 (12%) operations being carried out by a registrar aided by a consultant or vice versa. In the few cases in which a registrar was listed without a consultant in theatre ($n = 7$; 6%), all operating surgeons were identified as a senior registrar or specialist fellow and the availability of a consultant was confirmed in all cases.

Summary of surgery data

A summary of the surgery details from the surgical forms (see *Appendix 21*) is provided in *Table 11*. Details of the seven unexpected procedures listed are presented in *Box 2*.

Type of surgery/implant

Brief details of the actual types of implant provided to the 109 patients allocated to surgery were collected under four categories (nail, plate and screws, hemiarthroplasty and other). These are presented in *Table 12*.

The majority of fractures (82.6%) were fixed using a plate and screws, with the PHILOS plate being used in 66 (60.6%) cases. Hemiarthroplasty was used in 10 patients (9.2%); it is likely that cemented prostheses were used in at least eight of these.

Additional details on the implants, such as the specific type of PHILOS plate and number of screws used, were available from implant labels stuck, as requested, on the back of 12 surgery forms and written notes, recording various details such as the manufacturer's code for the device, on the back of a further 11 surgery forms. Two other forms noted the absence of implant labels to comply with this request. Although the general absence of implant labels was disappointing, the requests to centres for further key information to inform the cost analysis were generally successful (see *Chapter 7*).

TABLE 11 Treatment details: surgical methods used in the surgery group

Treatment	n (%)
Number randomised	125 (100)
Number receiving surgery ^a	109 (87.2)
Operation time ^b	
Time (minutes) in theatre, mean (SD)	113.14 (34.68)
Type of surgery (as listed)	
Nail	4 (3.7)
Plate and screws ^c	90 (82.6)
Hemiarthroplasty	10 (9.2)
Other	5 (4.6)
Unexpected procedures	7 (6.4)
Conversion ORIF to hemiarthroplasty	2 (1.8)
Rotator cuff repair	2 (1.8)
Other	3 (2.8)
Type of anaesthesia	
General ^d	107 (98.2)
Local ^e	21 (19.3)
Regional ^f	41 (37.6)
Other ^g	3 (2.8)
Antibiotics used	
Yes	100 (91.7)
No	1 (0.9)
Missing	8 (7.3)
Radiographic imaging	
Image intensifier	99 (90.8)
Films ^h	3 (2.8)

ORIF, open reduction and internal fixation.

a Data in table apply only to the patients who received surgery.

b Data missing for six patients.

c One patient also received a bone graft (Tutobone®; Wescott Medical, Chester-le-Street, UK) and another patient received two additional Kirschner (K) wires.

d The two patients not receiving general anaesthesia had regional anaesthesia.

e All patients received general anaesthesia as well.

f All but two patients received general anaesthesia as well.

g All three were in addition to either general or regional anaesthesia or both.

h Both image intensifier and films were used for one patient.

BOX 2 Unexpected surgical procedures

- Originally planned for ORIF – changed intraoperatively to Global hemiarthroplasty as graft tissue attachment did not allow fracture reduction.
- Attempted ORIF but reduction failed and therefore proceeded to hemiarthroplasty.
- Rotator cuff tear repair.
- Rotator cuff tear (pre-existing) – subacromial decompression and cuff closure.
- Fixation of a greater tuberosity fracture using FiberWire® sutures (Arthrex Ltd, Sheffield, UK).
- Femoral head used as a bone graft because of large defect.
- Fracture had united; take down thought hazardous and therefore no fixation was carried out.

ORIF, open reduction and internal fixation.

TABLE 12 Types of implant (*n* = 109)

Implant	<i>n</i> (%)	Implant types, <i>n</i>
Nail	4 (3.7)	Polarus® (Acumed, Andover, UK) (<i>n</i> = 3); Expert® (DePuy Synthes UK Ltd, Leeds, UK) (<i>n</i> = 1)
Plate and screws	90 (82.6)	PHILOS® (DePuy Synthes UK Ltd) (<i>n</i> = 66); AxSOS (Stryker UK Ltd, Newbury, UK) (<i>n</i> = 11); Acumed® plate (Acumed) (<i>n</i> = 3); NCB® humeral plate (Zimmer Inc., Warsaw, IN, USA) (<i>n</i> = 2); S3™ plate (DePuy Synthes UK Ltd) (<i>n</i> = 2); unnamed plates (<i>n</i> = 6)
Hemiarthroplasty	10 (9.2)	Global® (DePuy Synthes) (<i>n</i> = 6); Anatomical™ (Zimmer Inc.) ^a (<i>n</i> = 2); Biomet Bio-Modular® (Biomet UK Ltd, Swindon, UK) (<i>n</i> = 1); Epoca® (DePuy Synthes UK Ltd) (<i>n</i> = 1)
Other	5 (4.6)	Suture fixation (<i>n</i> = 1); titanium screws (<i>n</i> = 1); only screws (<i>n</i> = 1); fracture united – no fixation (<i>n</i> = 2)

NCB, non-contact bridging.

^a It is not known whether these were cemented or uncemented.

Interventions: non-surgical care

Brief details of the prescribed non-surgical treatment for these patients, exemplified by 'sling immobilisation', were collected using the treatment confirmation form. These details are summarised in Table 13. Of the 125 patients allocated to non-surgery, 78 (62.4%) were prescribed sling immobilisation, 35 (28.0%) were prescribed a collar and cuff; four (3.2%) were prescribed a polysling or equivalent and three (2.4%) were prescribed a hanging cast followed by a collar and cuff in one case and sling immobilisation in a second case. Data were missing for five patients. There was reference made to early mobilisation in six cases, to exercises in three cases and to physiotherapy in 18 cases. The prescribed duration of immobilisation was stated in four cases: 2 weeks, 2–3 weeks, 3 weeks and 6 weeks.

TABLE 13 Prescribed non-surgical care (sling immobilisation) received by non-surgery patients

Description	<i>n</i> (%)
Collar and cuff	35 (28.0)
Hanging cast	3 (2.4)
Polysling (or equivalent)	4 (3.2)
Sling	78 (62.4)
Missing	5 (4.0)

Information on sling immobilisation after surgery was not collected for individual patients. To gain further insights about this and the use of the ProFHER sling immobilisation information leaflet (see *Appendix 12*), we sent a survey to hospitals asking about current practice.

Findings from the questionnaire survey on sling immobilisation

All 35 participating sites completed the post-recruitment survey (see *Appendix 30*) on the provision of sling care information leaflets and sling immobilisation to patients eligible for participation in the ProFHER trial.

Use of the ProFHER trial sling immobilisation care leaflet

This was covered in the first page of the survey. The survey results show that 28 (85%) of the 33 centres that had identified eligible patients had routinely provided the ProFHER sling immobilisation care leaflet to eligible patients, whether or not they had consented to take part in the trial. Two sites 'never' used the leaflet, one of which provided an alternative. Three other sites 'rarely' used the leaflet. The reasons for not providing the leaflets were to not confuse the patient or complicate matters in terms of trial participation and because of administrative problems.

Inspection of the only alternative leaflet showed that it was a text-only leaflet providing general advice on sling use during conservative management of people with a proximal humeral fracture. It covered fewer items than the ProFHER leaflet; the only additional information being the advice that 'it may be more comfortable to wear the sling under your clothes for the first week'.

All 32 sites that had recruited patients into the trial routinely provided the ProFHER trial sling information leaflet. These results confirm the delivery of a standard of care, in particular for consenting patients, thereby ensuring comparability between the two randomised groups.

Type and duration of sling immobilisation typically provided to patients who would have been eligible for the ProFHER trial

This was covered in the second page of the survey. A distinction was made between the 32 centres that recruited patients and the three that did not.

Type of immobilisation for non-surgical patients

Twenty-two (69%) of the 32 centres that recruited patients reported routinely/typically using one method of immobilisation for non-surgically treated patients who would have been eligible for the ProFHER trial. The others reported either two (25%) or three (6%) methods. Comments provided by centres in which two or three devices were reported indicated that the choice of method was dependent on patient comfort in one centre, patient ability to cope in another centre and consultant preference in a third centre. *Table 14* summarises the type of immobilisation routinely/typically used for non-surgically treated patients in recruiting centres.

TABLE 14 Non-surgical treatments routinely/typically used by recruiting centres (*n* = 32)

Description	<i>n</i> (%)
Collar and cuff	14 (44)
Polysling (and equivalent)	8 (25)
Collar and cuff or polysling	7 (22)
Broad arm sling or polysling	1 (3)
Collar and cuff, broad arm sling or polysling	2 (6)

A collar and cuff was the main method used in 14 centres (44%) and was one of two or three methods used in a further nine centres (28%). A polysling or equivalent was the main method used in eight centres (25%) and one of two or three methods used in 10 centres (31%). Aside from those described under the generic label, we included the following devices in this category: cotton polysling, high arm sling (Chaneco, Chris Hanley & Partners, Northampton, UK), Lancaster sling (Promedics Orthopaedics Ltd, Port Glasgow, UK), body sling, shoulder immobiliser and Seaton sling. There was a general lack of information on the extent of immobilisation (i.e. use of the waist belt of the polysling). A broad arm sling was one of two or three methods of immobilisation in three centres. The occasional use of a hanging cast was noted in one centre. (The single methods described by the three centres that did not recruit patients were collar and cuff, polysling, and collar and hanging cast.) In principle, all of these slings work similarly, that is, to rest the arm comfortably adducted and internally rotated, with the arm by the side of the body, elbow flexed and the forearm resting in front of the body.

Type of immobilisation for post-surgical patients

Twenty-three (72%) of the 32 centres that recruited patients reported routinely/typically using one method of immobilisation for surgically treated patients who would have been eligible for the ProFHER trial. Six centres (19%) reported two methods, two (6%) reported three methods and one (3%) reported that immobilisation was a consultant-led decision based on the stability of fixation and extent of soft-tissue injury. Comments provided by centres in which two devices were reported indicated that the choice of immobilisation was dependent on the comfort and quality of fixation in one centre, on the presence or not of a cuff tear in one centre and on ADL and ability to cope in a third centre. *Table 15* summarises the types of immobilisation routinely/typically used post surgery in recruiting centres.

A polysling or equivalent was the main method in 17 centres (53%) and one of two or three methods in eight centres (25%). Aside from those described under the generic label, we included the following devices in this category: high arm sling (Chaneco), Lancaster sling, body sling, shoulder immobiliser, neoprene immobiliser and Seaton sling. There was a general lack of information on the extent of immobilisation (i.e. use of the body belt of the polysling), but the use of the immobiliser was specifically described in two cases. A collar and cuff was the main method in three centres (9%) and was one of two or three methods in a further seven centres (22%). A broad arm sling was the main method in two centres (6%) and was one of two or three methods of immobilisation in three other centres (9%). One centre described the use of a 'sling' only and another described policy rather than devices. (Information was missing from one of the three centres that did not recruit. Of the other two centres, the main method used was a polysling in one centre and a polysling without a waist belt in the other centre.)

In all, 14 centres used the same method or combination of methods for non-surgical and post-surgical immobilisation, as summarised in *Table 16*.

TABLE 15 Post-surgical immobilisation routinely/typically used by recruiting centres ($n = 32$)

Description	<i>n</i> (%)
Collar and cuff	3 (9)
Polysling	17 (53)
Broad arm sling	2 (6)
Not a defined sling	1 (3)
Collar and cuff or polysling	5 (16)
Broad arm sling or polysling	1 (3)
Collar and cuff, broad arm sling or polysling	2 (6)

TABLE 16 Recruiting centres listing the same methods for non-surgical and post-surgical immobilisation

Description	<i>n</i>
Collar and cuff	3
Polysling	7 ^a
Collar and cuff or polysling	3
Collar and cuff, broad arm sling or polysling	1
^a One centre used different but similar polyslings in the two groups.	

Duration of sling immobilisation for non-surgical patients

The recommended duration of arm immobilisation for non-surgically treated patients ranged from 2 to 6 weeks. Of the 32 recruiting centres, 29 (91%) recommended immobilisation of ≥ 3 weeks. This is consistent with our expectations and guidance. Commentary on this indicated that exercises often started during sling use: two centres referred to pendular exercises and six centres to exercises starting at around 2–3 weeks, with some reduction in the wearing of the sling. In five of these six centres, the treatment period was 6 weeks; in the other centre it was 4 weeks. In terms of sling immobilisation there was also reference made to fracture healing and dependency on type of fracture, pain, age and comorbidities and also, in four centres, to consultant preference.

Duration of sling immobilisation for post-surgical patients

The recommended duration of arm immobilisation after surgery was more variable, ranging from none (early mobilisation in one recruiting centre) to 8 weeks. Of the 32 recruiting sites, 23 (72%) recommended immobilisation of ≥ 3 weeks. This is consistent with our expectations but we gave no guidance on this. Commentary on this again indicated that exercises often started during sling use. In two centres, exercises started almost immediately, yet the recommended period of immobilisation was 6 weeks in both. In three other centres that reported immobilisation of 6 weeks, exercises/physiotherapy were reported to start at around 2–3 weeks. In one centre, sling use was for pain and comfort only, with no time given for the period of immobilisation. There was some mention of variation according to the method of surgery (four cases) but this was not in consistent directions for the two centres providing times (in one centre fixation involved longer immobilisation than hemiarthroplasty and in another centre the reverse was true). Reference was again made to fracture healing and dependency on the type of fracture and pain, with the additional consideration of the quality of surgical fixation in one centre. Four centres specifically reported that the decision depended on the surgeon.

In all, 19 (18 recruiting) centres reported the same recommended period of arm immobilisation for non-surgical treatment and post-surgical treatment.

Physiotherapy

Details of the physiotherapy provided to each treatment group, collected from the physiotherapy logs (see *Appendix 23*), are provided in *Table 17*. This demonstrates equivalence between the treatment groups in access to and implementation of physiotherapy, including advice given for home exercises and performance of home exercises by patients.

TABLE 17 Physiotherapy treatment

Description	Surgery (<i>n</i> = 125)	Not surgery (<i>n</i> = 125)
Number (%) receiving physiotherapy	118 (94.4)	117 (93.6)
Days to first session		
Mean (SD)	23.1 (24.07)	25.5 (18.60)
Median (min., max.)	16 (1, 120)	22 (-1, 122)
Duration of physiotherapy (days from first to last session)		
Mean (SD)	116.7 (75.75)	113.0 (67.46)
Median (min., max.)	111.5 (0, 510)	104 (0, 395)
Number of sessions		
Mean (SD)	9.6 (6.22)	9.6 (6.59)
Median (min., max.)	8 (1, 36)	8 (1, 43)
Allocated time per session (minutes)		
Mean (SD)	28.3 (9.57)	29.2 (10.49)
Median (min., max.)	30 (0, 60)	30 (0, 60)
Session details ^{a,b}		
Advice and/or education		
<i>n</i> (%)	115 (92.0)	113 (90.4)
Mean (SD)	7.1 (5.13)	7.3 (5.61)
Exercise		
<i>n</i> (%)	118 (94.4)	114 (91.2)
Mean (SD)	8.6 (5.37)	8.4 (6.08)
TENS		
<i>n</i> (%)	1 (0.8)	8 (6.4)
Mean (SD)	0.0 (0.09)	0.1 (0.34)
Soft-tissue techniques		
<i>n</i> (%)	49 (39.2)	49 (39.2)
Mean (SD)	1.3 (2.13)	1.9 (3.52)
Joint mobilisations		
<i>n</i> (%)	55 (44.0)	71 (56.8)
Mean (SD)	2.1 (3.11)	2.5 (3.40)
Stretching techniques		
<i>n</i> (%)	83 (66.4)	84 (67.2)
Mean (SD)	3.0 (3.16)	3.7 (4.11)

continued

TABLE 17 Physiotherapy treatment (continued)

Description	Surgery (n = 125)	Not surgery (n = 125)
Relaxation techniques		
n (%)	31 (24.8)	31 (24.8)
Mean (SD)	0.5 (1.11)	1.0 (2.95)
Hydrotherapy		
n (%)	5 (4.0)	10 (8.0)
Mean (SD)	0.2 (1.17)	0.5 (1.74)
Other, n (%)	22 (17.6)	22 (17.6)
Home exercise ^{a,b}		
Reported (yes or no)	112 (89.6)	106 (84.8)
Yes	109 (87.2)	103 (82.4)
Mean (SD)	6.9 (5.21)	7.0 (6.01)
No	38 (30.4)	44 (35.2)
Mean (SD)	0.6 (1.24)	0.5 (0.85)
Referral to other specialty, n (%) ^b	9 (7.2)	12 (9.6)
Occupational therapy, n (%)	1 (0.8)	2 (1.6)
Other, n (%)	8 (6.4)	10 (8.0)
Deviation from protocol, n (%)	0 (0.0)	0 (0.0)

Max., maximum; min., minimum.

a Frequency of patients if reported at least once in session logs.

b Means and SDs are based on the number of sessions for patients with available data.

Reasons for not starting physiotherapy

Based on the receipt of the physiotherapy logs and information from the physiotherapy end-of-treatment logs, 119 patients (95%) in the surgery group and 118 patients (94%) in the non-surgery group started physiotherapy. Thus, physiotherapy logs were missing for one patient in each group who started physiotherapy according to their end-of-treatment log. Of the six patients in the surgery group who were reported as not starting physiotherapy within the ProFHER trial, the reasons, when given, were non-attendance or patient unavailability in three cases and the separate provision of physiotherapy at the nursing home for one patient. Of the five patients in the non-surgery group who were reported as not starting physiotherapy, this was recorded as resulting from non-attendance or a change in health status or circumstances in four patients. Both types of physiotherapy forms were missing for two patients allocated to non-surgery. Overall, these data indicate that our efforts to ensure equivalent access to and delivery of physiotherapy were met for the two groups.

'Other' physiotherapy interventions

Twenty-two patients in each group received one or more 'other' physiotherapy interventions that were distinctly different from those listed in the physiotherapy logs or protocol. These do not include typical devices used for exercising, such as a Thera-Band (Patterson Medical Ltd, Sutton-in-Ashfield, UK) and pulleys, reports of which were excluded on screening of the data entries by two raters. 'Other' interventions reported were ice, heat, general mobility exercises, acupuncture, proprioceptive neuromuscular facilitations and ADL/functional training and advice.

Home exercises

Of the 118 patients in the surgery group for whom there was at least one completed physiotherapy log, 109 patients (92%) were recorded as having performed their home exercises in one or more logs. This applied to only slightly fewer patients, 103 (88%) of the 117 patients with completed physiotherapy logs, in the non-surgery group. The majority of patients were recorded as having performed their home exercises in most of their sessions. Just three patients in each group were recorded as never having carried out their home exercises.

Referrals to other specialties

When screening the entries for referrals to occupational therapy and other specialties in the physiotherapy logs, we discounted referrals to further physiotherapy for the shoulder, routine reviews with an orthopaedic consultant and referrals to a GP. Referrals to a GP were predominantly for patients to seek advice for pain medication and were considered as being for information only rather than a referral for a specific episode of care. Of the 18 referrals of individual patients to other specialties for 'problems', 12 were to an orthopaedic consultant and the others were single referrals to community care, neurology, radiology, a hand therapist, an oncologist and for physiotherapy aimed at balance and fall prevention.

Reasons for any substantial differences from the ProFHER trial physiotherapy protocol

The physiotherapy treatment log allowed a physiotherapist to record the reasons why the care delivered to a patient was substantially different from that in the physiotherapy protocol. Some variation from protocol was expected and permitted and, thus, referral to and description of adjustments to accommodate clinical need or instruction from the consultant were not considered as 'protocol violations'. One of us (HH) reviewed 486 treatment log comment entries and found no mention of electrotherapy, which would have been a 'protocol violation'. Reference was made to acupuncture, which was an accepted intervention. Lack of knowledge of/access to the physiotherapy protocol was identified from the comments made for two surgical group patients. In one case the physiotherapist providing treatment noted that they had no knowledge of the trial and in the second case a physiotherapist from an externally based clinic categorically stated that they did not know 'what the ProFHER physiotherapy protocol is'.

Provision of information leaflets illustrating home exercises

The trial protocol and trial manual indicated that physiotherapists involved in the treatment of trial patients should provide an information leaflet illustrating the home exercises. We did not assess compliance with this expectation. However, in discussions between the lead physiotherapy contact (LG) and the physiotherapists acting as the named contacts for physiotherapy at the participating hospital sites, we recorded that advice on home exercises was often adjusted to accommodate home exercises that were predominantly based on daily functional tasks (such as reaching up to the top of a cupboard). Written instructions for these were considered unhelpful and the emphasis was shifted to the monitoring and reinforcement of these and related functional activities at subsequent physiotherapy.

Completion of physiotherapy

Physiotherapy end-of-treatment logs were obtained for 245 patients. Similar reasons for formal discharge were provided for both treatment groups (*Table 18*). Although the time to formal discharge was on average 2 weeks later in the surgery group than in the non-surgery group, this is likely to have been administrative rather than reflecting outcome.

TABLE 18 Completion of physiotherapy

Description	Surgery (n = 125), n (%)	Not surgery (n = 125), n (%)
Treatment completed (formal discharge)	93 (74.4)	96 (76.8)
Independent shoulder function achieved	66 (52.8)	72 (57.6)
No improvement noted over several sessions	19 (15.2)	18 (14.4)
Other	8 (6.4)	6 (4.8)
Treatment not completed (formal discharge)	23 (18.4)	22 (17.6)
Patient stopped attending	17 (13.6)	14 (11.2)
Another problem intervened	3 (2.4)	2 (1.6)
Other	3 (2.4)	5 (4.0)
Unknown	0 (0)	1 (0.8)
Treatment never started (patient never attended)	6 (4.8)	5 (4.0)
Days to discharge since start of therapy, mean (SD)		
Overall	137.5 (83.6)	123.0 (71.6)
Formal discharge (treatment completed)	141.3 (80.5)	127.5 (68.9)
Formal discharge (treatment not completed)	121.7 (96.0)	102.7 (81.2)

Return of data collection forms

Patient questionnaires

In total, follow-up patient questionnaires were returned for 225 (90%), 232 (93%), 226 (90%) and 218 (87%) of the 250 recruited patients at 3, 6, 12 and 24 months' follow-up respectively. Thus, we achieved a response rate above the 80% target stated in our protocol. These statistics include the abridged questionnaire data (including the OSS) collected by telephone as a last resort in 19 instances: nine (four surgery vs. five non-surgery) at 6 months, seven (three surgery vs. four non-surgery) at 12 months and three (two surgery vs. one non-surgery) at 24 months. The return of questionnaires by treatment group is shown in *Table 19*.

Patient questionnaires were dispatched promptly from the YTU on the date that the follow-up assessment was due; a letter informing patients to expect the questionnaire was sent a fortnight beforehand. The number of days from questionnaire dispatch by the YTU to the completion of the questionnaire as recorded by the patient ('Days to completion') and the number of days from questionnaire dispatch by the YTU to the returned questionnaire being logged by the YTU ('Days to return') are presented in *Table 20*. Medians and interquartile statistics are presented given the right-skewed distribution, as the majority of

TABLE 19 Return of patient questionnaires

Questionnaire	Surgery, n (%)	Not surgery, n (%)
Baseline	125	125
3 months' follow-up	114 (91)	111 (89)
6 months' follow-up	113 (90)	119 (95)
12 months' follow-up	111 (89)	115 (92)
24 months' follow-up	109 (87)	109 (87)

TABLE 20 Days from due follow-up date to completion (date recorded on the patient questionnaire) and return (logged by the YTU) of the questionnaires

Follow-up	Surgery		Not surgery	
	Days to completion ^a	Days to return	Days to completion ^a	Days to return
3 months, <i>n</i> , median (IQR)	93, 7 (4–16)	114, 12 (8–22)	99, 8 (4–12)	111, 11 (8–18)
6 months, <i>n</i> , median (IQR)	97, 6 (3–13)	113, 11 (8–22)	106, 7 (4–12)	119, 12 (8–18)
12 months, <i>n</i> , median (IQR)	93, 6 (4–11)	111, 12 (8–18)	101, 7 (4–12)	115, 13 (8–20)
24 months, <i>n</i> , median (IQR)	89, 7 (4–15)	109, 13 (10–21)	95, 7 (4–13)	109, 11 (8–19)

^a The lower numbers of patients in the days–completion columns than in the days to return columns reflect the fact that some patients at each time period did not record the date on which they completed the questionnaire.

patients completed and returned their forms promptly. As expected, the time to completion was shorter than the time to return. There was no obvious difference between the two groups in the time to completion or the time to return.

Return of hospital data collection forms

The return of the hospital forms is summarised in *Table 21*. Treatment confirmation forms, which were collected 1 month after treatment, were collected for all patients. This provided confirmation that the surgical forms for the 109 participants who had their allocated surgery had been returned, as had the inpatient episode forms for all 130 participants who had been admitted to hospital. Overall, 235 of 250 (94.0%) physiotherapy treatment logs and 245 of 250 (98.0%) physiotherapy end-of-treatment forms were returned. In addition, 249 of 250 (99.6%) 1-year follow-up forms and 234 of 250 (93.6%) 2-year follow-up forms were returned.

TABLE 21 Return of hospital forms

Hospital form	Surgery (<i>n</i> = 125), <i>n</i> (%)	Not surgery (<i>n</i> = 125), <i>n</i> (%)
Treatment confirmation form	125 (100)	125 (100)
Surgical form	109 (100 ^a)	[6 ^b]
Inpatient episode form	109 (100 ^a) [2 ^b]	22 (100 ^a) [1 ^b]
Physiotherapy treatment log (any received for patient)	118 (99 ^c)	117 (98 ^c)
Physiotherapy end-of-treatment log	122 (98)	123 (98)
1-year follow-up form	124 (99)	125 (100)
2-year follow-up form ^d	113 (90)	121 (97)

^a Percentage based on the number expected from reports in the treatment confirmation forms.

^b Additional forms received.

^c Percentage based on number expected from accounts in the physiotherapy end-of-treatment logs (forms not received for one surgery and two not surgery patients who started physiotherapy; two not surgery physiotherapy end-of-treatment logs missing).

^d Forms were not requested for those known to have died within 1 year.

Other returns

Baseline radiographs

Anonymised copies of baseline radiographs were obtained for all 250 patients. Although provision had been allowed for copies of radiographs to be sent as films, all copies of baseline radiographs were returned in digital format. This was consistent with the completion in December 2007 of the national roll-out of PACS to all acute trusts across England as part of the NHS Connecting for Health initiative.^{67,68} There were, however, greater than anticipated difficulties in obtaining copies of images that met the criteria for blinding and transferability (electronic portability). Hospitals used different PACS packages and different settings for permission of access for saving images as JPEG files. All images were under the requested 300 dpi resolution. When checked, the majority of images were at 96 dpi resolution, which appears to be the accepted resolution for images that would be viewed by surgeons in everyday practice. Ultimately, JPEG files were available for all baseline radiographs; in 26 sets of radiographs these were screen shots because of unresolved difficulties with saving images from the various PACS packages used.

Monitoring of the radiographs confirmed that all fractures met the main inclusion criterion that fractures should involve the surgical neck. No site was contacted with specific concerns relating to this issue. Aside from basic checks regarding blinding and requests for additional views when two views were not received, we did not ask for 'better quality radiographs' because it seemed safe to assume that we had been sent the best quality baseline radiographs available.

Adverse events

The 88 adverse event (reporting) forms (see *Appendix 27*) returned reported adverse events for 61 patients occurring within the 2-year follow-up period. Seventy-eight SAEs were reported for 56 patients and 10 non-SAEs were reported for nine patients. As judged by the sites, five SAEs met the criteria for informing the MREC. On checks made by two reviewers (HH and AR) of these and all SAEs, and in consultation with the DMEC, three of these and one other adverse event were reported separately to the MREC. No further action was required for any of these four events. For all 1-month review of adverse event forms (see *Appendix 28*), both reviewers agreed that there was no cause for further action to be taken by sites to address the adverse events and progress made by the patients involved.

Patient withdrawals

Two patients allocated to surgery explicitly withdrew from their treatment but agreed to complete the patient questionnaires. Six patients withdrew from completing the questionnaires at different stages of follow-up; all six confirmed that we could continue to collect data from their hospital records. No patient withdrew entirely from the trial.

Chapter 5 Primary outcome (Oxford Shoulder Score) analysis

The OSS [0 (worst score) to 48 (best score)] was assessed at 6, 12 and 24 months. When a maximum of two of the 12 OSS items were missing, these were replaced with the mean of the remaining items.⁴¹ An effect size of 0.4 (assuming a SD of 12) was sought between treatment groups; this is equivalent to an OSS difference of approximately 5 points.

Missing data

The multilevel model analysis used assumed any missing data to be missing at random (MAR). The extent of and reasons for missing OSS data at each time point are detailed in *Table 22*. Complete response is defined here as sufficient data (maximum of two missing OSS items) to allow the calculation of the OSS total score.

Overall, 209 patients (84%) had complete OSS responses at all follow-up time points (surgery: 103 patients; not surgery: 106 patients), whereas 41 patients (16%) did not respond or only partially responded at one or more time points (surgery: 22 patients; not surgery: 19 patients). The latter group are defined as non-responders in the following tables and analysis.

Demographic and fracture characteristics at baseline for responders and non-responders are presented in *Tables 23* and *24* respectively. Each baseline variable was entered into a logistic regression model predicting non-response, none of which reached statistical significance ($p < 0.10$). Consequently, no additional variables were included in the primary analysis.

TABLE 22 Missing primary outcome data over time

Follow-up	Response type	Surgery ($n = 125$), n (%)	Not surgery ($n = 125$), n (%)
6 months	Complete response	113 (90.4)	119 (95.2)
	No or partial response	12 (9.6)	6 (4.8)
	No response	8 (6.4)	4 (3.2)
	Withdrawn from questionnaires	1 (0.8)	1 (0.8)
	Died	3 (2.4)	1 (0.8)
12 months	Complete response	111 (88.8)	114 (91.2)
	No or partial response	14 (11.2)	11 (8.8)
	Insufficient data for calculation of OSS	0 (0.0)	1 (0.8)
	No response	4 (3.2)	6 (4.8)
	Withdrawn from questionnaires	3 (2.4)	2 (1.6)
	Died	7 (5.6)	2 (1.6)
24 months	Complete response	106 (84.8)	109 (87.2)
	No or partial response	19 (15.2)	16 (12.8)
	Insufficient data for calculation of OSS	3 (2.4)	0 (0.0)
	No response	3 (2.4)	9 (7.2)
	Withdrawn from questionnaires	4 (3.2)	2 (1.6)
	Died	9 (7.2)	5 (4.0)

TABLE 23 Predictors of non-response: demographics

Characteristic	Complete response (<i>n</i> = 209)	No/partial response (<i>n</i> = 41)	<i>p</i> -value (prediction of non-response)
Gender, <i>n</i> (%)			0.836
Male	49 (23.4)	9 (22.0)	
Female	160 (76.6)	32 (78.0)	
Age (years)			0.763
<i>n</i>	209	41	
Mean (SD)	66.1 (11.50)	66.5 (14.10)	
Median (min., max.)	66.8 (31.3, 89.0)	67.4 (24.6, 92.0)	
Ethnicity, <i>n</i> (%)			NA
White	208 (99.5)	41 (100.0)	
Black	1 (0.5)	0 (0.0)	
Education, <i>n</i> (%)			0.233
No formal qualifications	107 (51.2)	27 (65.9)	
Some qualifications/no degree	79 (37.8)	11 (26.8)	
Degree or higher	23 (11.0)	3 (7.3)	
Employment, <i>n</i> (%)			0.436
Part-time	19 (9.1)	0 (0.0)	
Full-time	32 (15.3)	7 (17.1)	
Self-employed	4 (1.9)	0 (0.0)	
Retired	133 (63.6)	27 (65.9)	
Not employed but seeking work	4 (1.9)	0 (0.0)	
Other	15 (7.2)	6 (14.6)	
Missing	2 (1.0)	1 (2.4)	
Diabetes, <i>n</i> (%)			0.336
Yes	24 (11.5)	7 (17.1)	
No	183 (87.6)	34 (82.9)	
Missing	2 (1.0)	0 (0.0)	
Smoking status, <i>n</i> (%)			0.188
Yes	50 (23.9)	14 (34.1)	
No	151 (72.2)	26 (63.4)	
Missing	8 (3.8)	1 (2.4)	
Steroid use, <i>n</i> (%)			0.424
Yes	12 (5.7)	1 (2.4)	
No	196 (93.8)	38 (92.7)	
Missing	1 (0.5)	2 (4.9)	
Health status (EQ-5D index)			0.136
<i>n</i>	204	40	
Mean (SD)	0.39 (0.36)	0.48 (0.40)	
Median (min., max.)	0.29 (−0.35, 1)	0.62 (−0.36, 1)	
Max., maximum; min., minimum; NA, not applicable.			

TABLE 24 Predictors of non-response: fracture data

Characteristic	Complete response (<i>n</i> = 209)	No/partial response (<i>n</i> = 41)	<i>p</i> -value (prediction of non-response)
Time since injury (days)			0.784
<i>n</i>	209	41	
Mean (SD)	5.7 (4.89)	5.9 (4.92)	
Median (min., max.)	4 (0, 21)	4 (1, 18)	
Affected shoulder, <i>n</i> (%)			0.394
Left	102 (48.8)	23 (56.1)	
Right	107 (51.2)	18 (43.9)	
Tuberosity involvement, <i>n</i> (%)			0.332
Tuberosity not involved or missing	45 (21.5)	12 (29.3)	
Greater tuberosity	104 (49.8)	15 (36.6)	
Lesser tuberosity	7 (3.3)	3 (7.3)	
Greater and lesser tuberosity	53 (25.4)	11 (26.8)	
Previous fractures, <i>n</i> (%)			0.674
Yes	54 (25.8)	12 (29.3)	
No	153 (73.2)	29 (70.7)	
Missing	2 (1.0)	0 (0.0)	
Previous surgery for fractures, <i>n</i> (%)			0.391
Yes	15 (7.2)	5 (12.2)	
No	37 (17.7)	7 (17.1)	
Missing	2 (1.0)	0 (0.0)	
No previous fractures	155 (74.2)	29 (70.7)	
Shoulder dominance, <i>n</i> (%)			0.948
Yes	107 (51.2)	21 (51.2)	
No	99 (47.4)	19 (46.3)	
Missing	3 (1.4)	1 (2.4)	
Injury mechanism, <i>n</i> (%)			0.510
Fall/trip from standing height or less	156 (74.6)	30 (73.2)	
Fall downstairs/steps or from a height	30 (14.4)	5 (12.2)	
Other	18 (8.6)	6 (14.6)	
Missing	5 (2.4)	0 (0.0)	
Max., maximum; min., minimum.			

Table 25 provides the descriptive OSS total scores for complete responders and those non-responders for whom data were available at some of the time points. Complete responders had better OSS outcomes than non-responders at all time points. OSS scores were entered into a logistic regression model predicting non-response and the observed differences were statistically significant at 12 months' follow-up ($p = 0.025$) but not at 6 or 24 months' follow-up. Although this is a clear indication that data were not MAR, this is not anticipated to be a substantial problem given the relatively small number of non-responders.

TABLE 25 Predictors of non-response: OSS total score

Characteristic	Complete response (<i>n</i> = 209)	No/partial response (<i>n</i> = 41)	<i>p</i> -value (prediction of non-response)
OSS at 6 months			0.193
<i>n</i>	209	23	
Mean (SD)	34.83 (10.38)	31.78 (12.39)	
Median (min., max.)	38 (3, 48)	36 (3, 48)	
OSS at 12 months			0.025
<i>n</i>	209	16	
Mean (SD)	37.13 (10.55)	30.63 (12.59)	
Median (min., max.)	41 (1, 48)	30.5 (4, 46)	
OSS at 24 months			0.167
<i>n</i>	209	6	
Mean (SD)	38.49 (10.41)	32.33 (10.27)	
Median (min., max.)	42 (1, 48)	30.5 (23, 46)	

Max., maximum; min., minimum.

Descriptive Oxford Shoulder Score statistics

Descriptive OSS statistics at 6, 12 and 24 months are presented in *Table 26* and illustrated in *Figure 8*. Mean scores improved over time for both treatment groups. At 6 months, the surgery group showed better scores than the non-surgery group (a mean difference of 3.0 points), although the 95% confidence interval (CIs) overlapped. This difference was smaller than the sought effect size of 5 points. Mean scores at 12 and 24 months were very similar between the groups.

TABLE 26 Unadjusted descriptive OSS statistics over time by treatment group

Follow-up	Surgery (<i>n</i> = 125)	Not surgery (<i>n</i> = 125)	Total (<i>n</i> = 250)
6 months			
<i>n</i> (%)	113 (90.4)	119 (95.2)	232 (92.8)
Mean (SD) score	36.07 (9.99)	33.07 (11.00)	34.53 (10.61)
Median (min., max.) score	39 (11, 48)	35 (3, 48)	37 (3, 48)
12 months			
<i>n</i> (%)	111 (88.8)	114 (91.2)	225 (90.0)
Mean (SD) score	36.89 (10.78)	36.45 (10.86)	36.67 (10.80)
Median (min., max.) score	41 (1, 48)	41 (4, 48)	41 (1, 48)
24 months			
<i>n</i> (%)	106 (84.8)	109 (87.2)	215 (86.0)
Mean (SD) score	38.25 (9.91)	38.39 (10.96)	38.32 (10.43)
Median (min., max.) score	42 (16, 48)	42 (1, 48)	42 (1, 48)

Max., maximum; min., minimum.

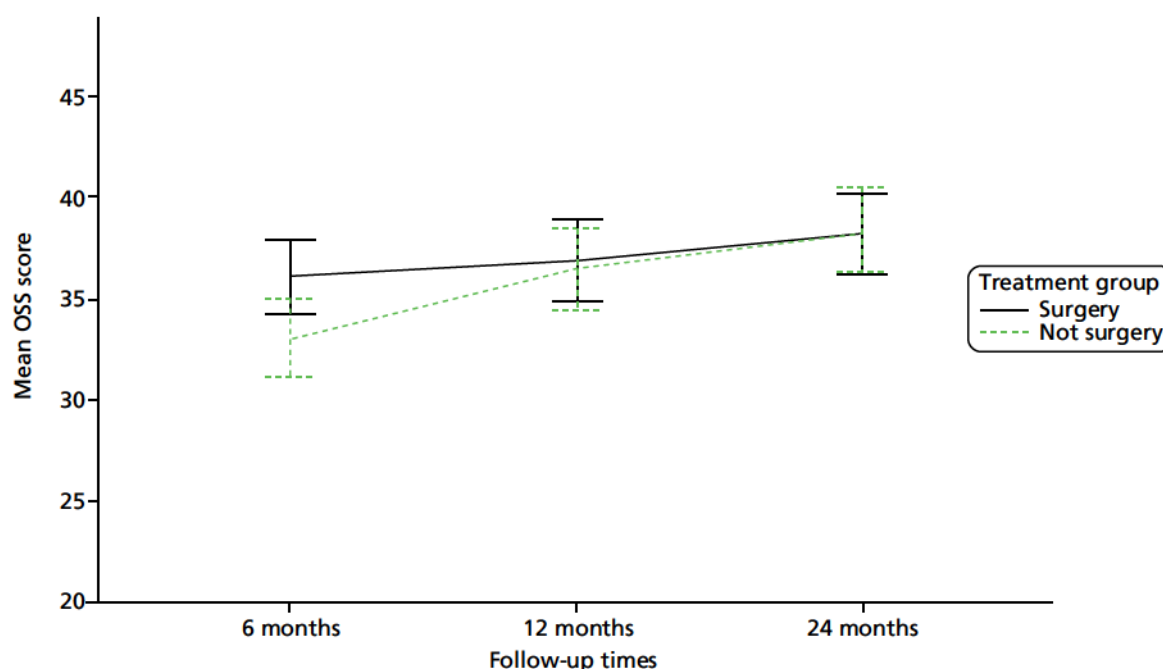


FIGURE 8 Mean OSS scores (with 95% confidence intervals) over time by treatment group.

As suggested by the discrepancy between the means and the medians, the distribution of OSS scores was found to be left skewed (*Figure 9*), that is, a greater number of participants reported higher (better) shoulder function.

Primary analysis

The primary outcome was analysed using a multilevel model. Patients were treated as a random effect and time (6, 12 and 24 months) was nested within patients. The effect of treatment group was assessed while adjusting for time, group \times time interaction, baseline EQ-5D index, gender, age group and tuberosity involvement at baseline. As no baseline characteristics were found to predict non-response, no further covariates were added.

Different covariance structures were applied to this model. An unstructured pattern that models all variances and covariances separately was judged to be the most intuitive. The more restrictive patterns of independent and exchangeable structures were tried and resulted in a worse model fit. Therefore, the unstructured covariance pattern was selected.

Diagnostics of model fit revealed non-normal standard residuals that were not uniform against predicted model values. This was likely to be a result of the left-skewed distribution of OSS scores. Therefore, all analyses were carried out on squared OSS values, which reduced skew and resulted in acceptable fit diagnostics. All group means and differences reported here were back transformed to the original OSS scores.

Adjusted OSS means and group differences for the model as specified above are presented in *Table 27*. The primary analysis (and subgroup analyses) included data from 231 patients (114 surgery, 117 not surgery) with a valid OSS score for at least one follow-up time point and complete baseline covariates. The analysis showed no significant differences between treatment groups at any time point, although the difference of 2.25 score points in favour of the surgery group at 6 months approached significance ($p = 0.058$). There was no overall effect of treatment group (difference of 0.75 score points in favour of the surgery group; $p = 0.479$).

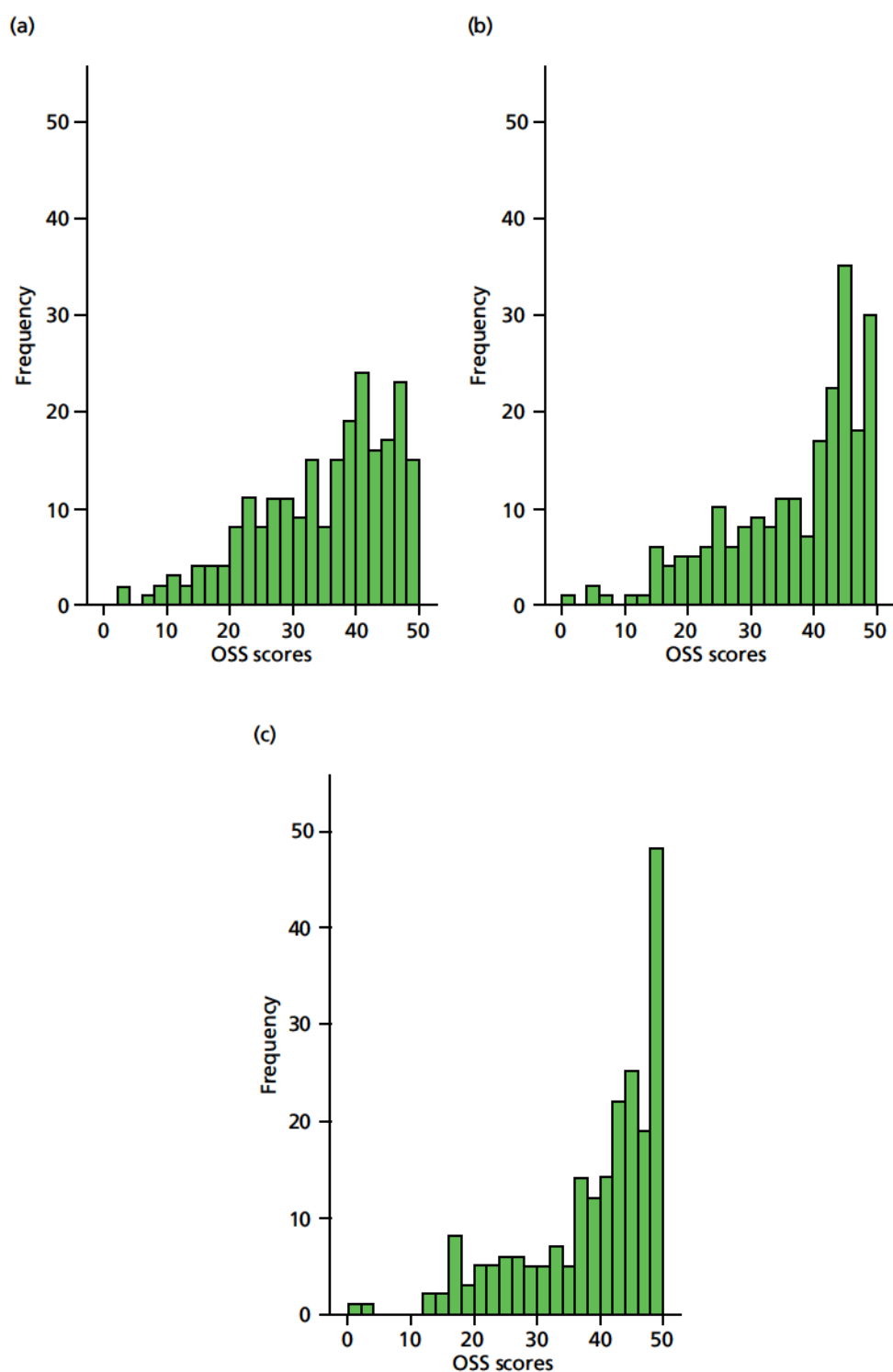


FIGURE 9 Distribution of OSS scores over time. (a) 6 months; (b) 12 months; and (c) 24 months.

TABLE 27 Difference in mean OSS scores over time by treatment group: primary analysis using a multilevel model^{a,b}

Follow-up	Surgery (<i>n</i> = 125), mean (95% CI), <i>n</i>	Not surgery (<i>n</i> = 125), mean (95% CI), <i>n</i>	Difference (95% CI), standard error ^c	<i>p</i> -value
Overall	39.07 (37.30 to 40.76), 114	38.32 (36.57 to 39.99), 117	0.75 (−1.33 to 2.84), 1.0638	0.479
6 months	37.84 (35.93 to 39.65), 111	35.59 (33.62 to 37.45), 115	2.25 (−0.07 to 4.57), 1.1837	0.058
12 months	39.23 (37.38 to 40.99), 109	38.80 (36.99 to 40.53), 110	0.42 (−1.78 to 2.63), 1.1250	0.706
24 months	40.11 (38.24 to 41.90), 104	40.40 (38.59 to 42.13), 106	−0.29 (−2.53 to 1.95), 1.1429	0.800

a Multilevel model adjusted for treatment group, time (6, 12 and 24 months), group × time interaction, baseline EQ-5D index, gender, age group (< 65 years/≥ 65 years) and tuberosity involvement at baseline (yes/no).

b The model included data from 114 surgery and 117 not surgery patients with valid OSS data for at least one follow-up time point and complete baseline covariates. The *n* values at the individual time points, which are presented for information only, are the numbers with valid OSS data and complete baseline covariates at that time point.

c Standard errors are the more appropriate measure of variability for these data than SDs.

Subgroup analyses

Age

The first subgroup analysis was undertaken to test the hypothesis that younger patients (aged < 65 years) would benefit more from surgical treatment than older patients. The relationship between age and treatment group in terms of OSS is illustrated in *Table 28* and *Figure 10*. Older patients generally had lower OSS scores (including poorer shoulder functioning), and scores in the surgery group were higher than those in the non-surgery group at all time points for this age group, although the CIs overlapped in each case. There was a trend for younger patients in the non-surgery group to improve more rapidly than those in the surgery group and older patients.

The same multilevel model as for the primary analysis was applied but an additional treatment group × age group interaction was included. Adjusted OSS means and group differences for this model are presented in *Table 29*. Similar to the primary analysis, there was no significant overall effect of treatment group over the 2-year follow-up period and no significant effect of treatment group at individual time points. A 2-log likelihood comparison with the primary analysis model was not significant [$\chi^2(1) = 1.24$, $p = 0.265$]; therefore, inclusion of the interaction did not significantly improve the model.

Fracture type

The second subgroup analysis was undertaken to test the hypothesis that patients with more complex fractures would benefit more from surgical treatment than patients with less complex fractures. Baseline radiographs were initially assessed by surgeons at the time of assessing patient eligibility in terms of tuberosity involvement (primary subgroup analysis) and were later categorised by the two independent surgeons using the Neer classification (sensitivity subgroup analysis).

Tuberosity involvement at baseline

Surgeons indicated at baseline whether a fracture involved the greater tuberosity, the lesser tuberosity or both in addition to the surgical neck. This information was summarised as tuberosity involvement (yes or no). The relationship between tuberosity involvement and treatment group in terms of OSS scores is illustrated in *Table 30* and *Figure 11*. Patients with fractures that involved one or both tuberosities generally had lower OSS scores (poorer shoulder functioning). Again, the scores in the surgery group were noticeably higher than those in the non-surgery group at 6 months, although the CIs overlapped. The differences between the two groups were greater for patients with fractures that involved neither tuberosity than for those with fractures with tuberosity involvement.

TABLE 28 Descriptive OSS statistics over time by treatment group and age group

Follow-up	Surgery (<i>n</i> = 125)	Not surgery (<i>n</i> = 125)	Total (<i>n</i> = 250)
Age: < 65 years (<i>n</i> = 108)			
<i>6 months</i>			
<i>n</i>	46	53	99
Mean (SD)	35.5 (10.87)	33.9 (10.11)	34.7 (10.45)
Median (min., max.)	39 (11, 48)	35 (3, 48)	38 (3, 48)
<i>12 months</i>			
<i>n</i>	48	53	101
Mean (SD)	37.0 (11.6)	37.8 (10.53)	37.4 (11.01)
Median (min., max.)	42 (5, 48)	42 (4, 48)	42 (4, 48)
<i>24 months</i>			
<i>n</i>	45	50	95
Mean (SD)	38.7 (9.95)	41.0 (8.36)	39.9 (9.17)
Median (min., max.)	43 (16, 48)	43 (14, 48)	43 (14, 48)
Age: ≥ 65 years (<i>n</i> = 142)			
<i>6 months</i>			
<i>n</i>	67	66	133
Mean (SD)	36.4 (9.40)	32.4 (11.72)	34.4 (10.76)
Median (min., max.)	39 (14, 48)	34.5 (3, 48)	37 (3, 48)
<i>12 months</i>			
<i>n</i>	63	61	124
Mean (SD)	36.8 (10.20)	35.2 (11.09)	36.0 (10.63)
Median (min., max.)	40 (1, 48)	39 (7, 48)	40 (1, 48)
<i>24 months</i>			
<i>n</i>	61	59	120
Mean (SD)	37.9 (9.94)	36.2 (12.41)	37.1 (11.21)
Median (min., max.)	41 (16, 48)	39 (1, 48)	40.5 (1, 48)
Max., maximum; min., minimum.			

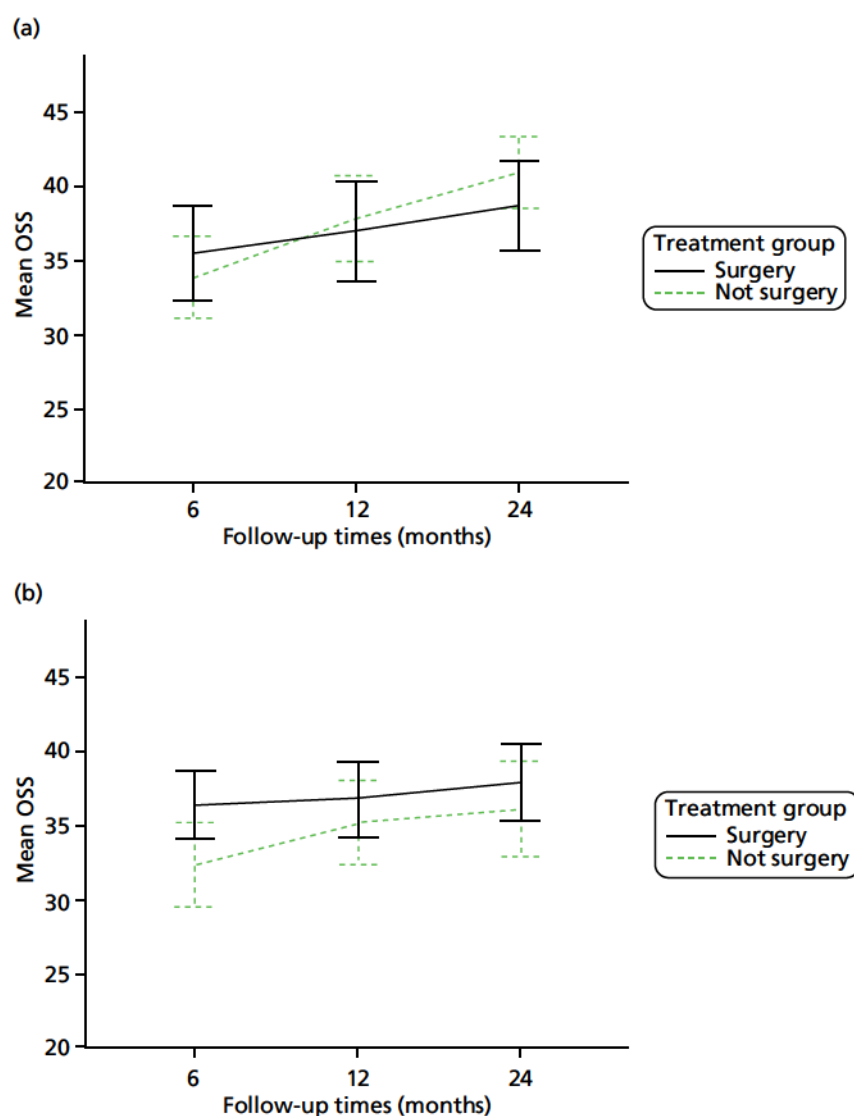


FIGURE 10 Mean OSS scores (with 95% CIs) over time by treatment group and age group. (a) < 65 years and (b) ≥ 65 years.

TABLE 29 Difference in mean OSS scores over time by treatment group: subgroup analysis 1 – age (< 65 years/≥ 65 years)^a

Follow-up	Surgery (<i>n</i> = 125; <i>n</i> = 114 in analysis), mean (95% CI)	Not surgery (<i>n</i> = 125; <i>n</i> = 117 in analysis), mean (95% CI)	Difference (95% CI)	<i>p</i> -value
Overall	38.98 (37.21 to 40.68)	38.40 (36.66 to 40.07)	0.58 (–1.52 to 2.68)	0.590
6 months	37.74 (35.83 to 39.56)	35.69 (33.72 to 37.55)	2.06 (–0.28 to 4.40)	0.085
12 months	39.14 (37.29 to 40.91)	38.89 (37.08 to 40.62)	0.26 (–1.96 to 2.47)	0.822
24 months	40.03 (38.15 to 41.82)	40.48 (38.68 to 42.21)	–0.46 (–2.71 to 1.79)	0.690

^a Multilevel model adjusted for treatment group, time (6, 12 and 24 months), group × time interaction, baseline EQ-5D index, gender, age (< 65 years/≥ 65 years), group × age interaction and tuberosity involvement at baseline (yes/no).

TABLE 30 Descriptive OSS statistics over time by treatment group and tuberosity involvement

Follow-up	Surgery (<i>n</i> = 125)	Not surgery (<i>n</i> = 125)	Total (<i>n</i> = 250)
<i>Neither tuberosity involved (n = 57)</i>			
<i>6 months</i>			
<i>n</i>	24	29	53
Mean (SD)	37.6 (9.81)	33.6 (11.89)	35.4 (11.09)
Median (min., max.)	41.5 (18, 48)	36 (3, 48)	38 (3, 48)
<i>12 months</i>			
<i>n</i>	22	28	50
Mean (SD)	38.7 (10.58)	36.6 (11.92)	37.5 (11.29)
Median (min., max.)	44 (14, 48)	43 (7, 48)	43 (7, 48)
<i>24 months</i>			
<i>n</i>	21	26	47
Mean (SD)	39.0 (9.86)	40.0 (12.34)	39.6 (11.19)
Median (min., max.)	43 (16, 48)	45.5 (1, 48)	44 (1, 48)
<i>Greater and/or lesser tuberosity involved (n = 193)</i>			
<i>6 months</i>			
<i>n</i>	89	90	179
Mean (SD)	35.7 (10.05)	32.9 (10.78)	34.3 (10.48)
Median (min., max.)	39 (11, 48)	33.5 (3, 48)	37 (3, 48)
<i>12 months</i>			
<i>n</i>	89	86	175
Mean (SD)	36.4 (10.84)	36.4 (10.57)	36.4 (10.68)
Median (min., max.)	40 (1, 48)	40 (4, 48)	40 (1, 48)
<i>24 months</i>			
<i>n</i>	85	83	168
Mean (SD)	38.1 (9.97)	37.9 (10.53)	38.0 (10.23)
Median (min., max.)	41 (16, 48)	42 (3, 48)	41.5 (3, 48)
Max., maximum; min., minimum.			

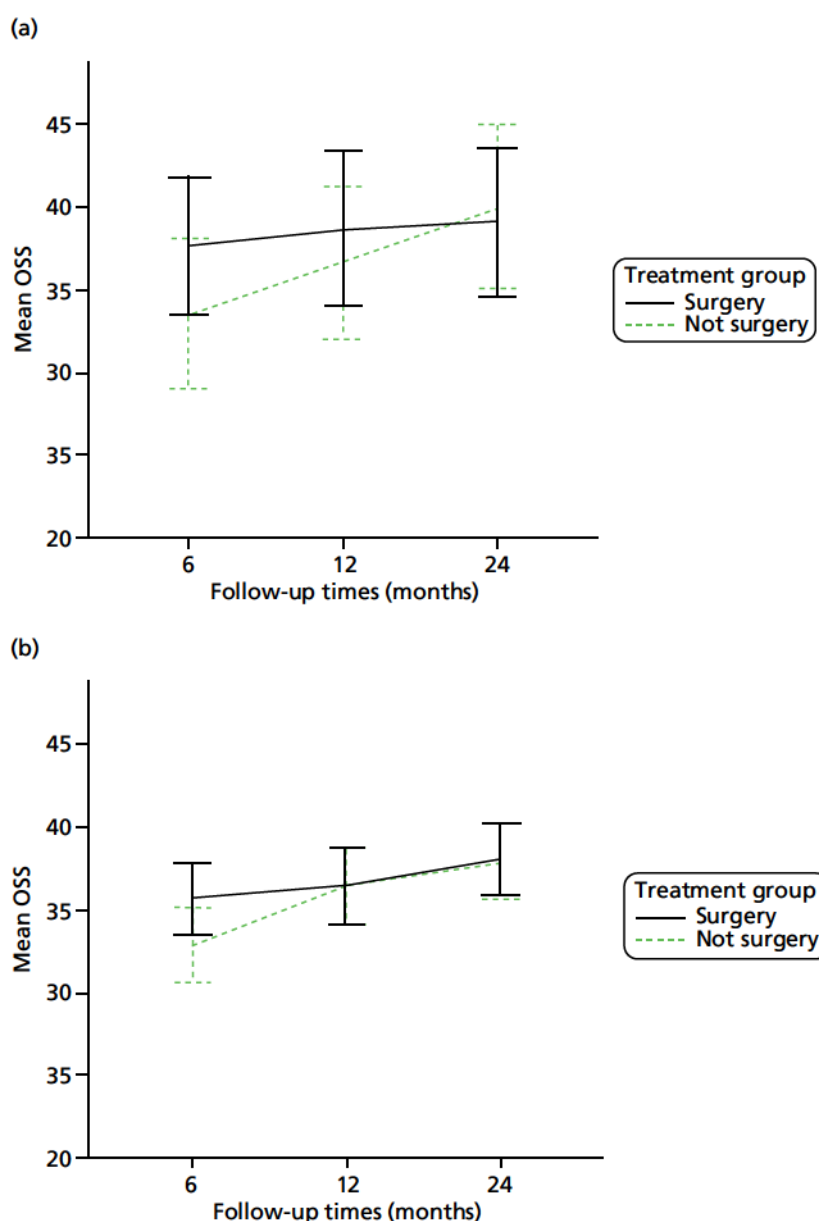


FIGURE 11 Mean OSS scores (with 95% CIs) over time by treatment group and tuberosity involvement. (a) Neither tuberosity involved and (b) greater and/or lesser tuberosity involved.

The same multilevel model as for the primary analysis was applied but an additional treatment group by tuberosity involvement interaction was included. Adjusted OSS means and group differences for this model are presented in *Table 31*. Similar to the primary analysis, there was no significant overall effect of treatment group over the 2-year follow-up period and no significant effect of treatment group at individual time points. A 2-log likelihood comparison with the primary analysis model was not significant [$\chi^2(1) < 0.01$, $p = 0.954$]; therefore, inclusion of the interaction did not significantly improve the model.

Independent Neer classification (sensitivity analysis)

The Neer classification categories were grouped into one- or two-part fractures and three- or four-part fractures. The relationship between Neer grouping and treatment group in terms of OSS scores is illustrated in *Table 32* and *Figure 12*. Patients with Neer three- or four-part fractures had lower OSS scores (poorer shoulder functioning), and scores in the surgery group were higher than those in the non-surgery group at all time points, although CIs overlapped. Patients with Neer one- or two-part fractures seemed to improve shoulder functioning at a slower rate over time in the surgery group than in the non-surgery group.

TABLE 31 Difference in mean OSS scores over time by treatment group: subgroup analysis 2 – tuberosity involvement (yes/no)^a

Follow-up	Surgery (<i>n</i> = 125; <i>n</i> = 114 in analysis), mean (95% CI)	Not surgery (<i>n</i> = 125; <i>n</i> = 117 in analysis), mean (95% CI)	Difference (95% CI)	<i>p</i> -value
Overall	39.05 (37.10 to 40.90)	38.33 (36.48 to 40.10)	0.71 (–1.79 to 3.22)	0.577
6 months	37.81 (35.73 to 39.79)	35.61 (33.53 to 37.57)	2.20 (–0.54 to 4.94)	0.115
12 months	39.20 (37.18 to 41.12)	38.82 (36.91 to 40.64)	0.38 (–2.22 to 2.98)	0.772
24 months	40.09 (38.05 to 42.02)	40.42 (38.52 to 42.23)	–0.33 (–2.94 to 2.28)	0.804

a Multilevel model adjusted for treatment group, time (6, 12 and 24 months), group × time interaction, baseline EQ-5D index, gender, age (< 65 years/≥ 65 years), tuberosity involvement at baseline (yes/no) and group × tuberosity involvement interaction.

TABLE 32 Descriptive OSS statistics over time by treatment group and Neer classification

Follow-up	Surgery (<i>n</i> = 125)	Not surgery (<i>n</i> = 125)	Total (<i>n</i> = 250)
Neer one- and two-part fractures (<i>n</i> = 146)			
<i>6 months</i>			
<i>n</i>	69	67	136
Mean (SD)	37.4 (9.69)	34.6 (10.59)	36.0 (10.20)
Median (min., max.)	41 (12, 48)	37 (3, 48)	39 (3, 48)
<i>12 months</i>			
<i>n</i>	66	62	128
Mean (SD)	38.2 (10.51)	38.4 (10.62)	38.3 (10.52)
Median (min., max.)	42.5 (5, 48)	43 (4, 48)	43 (4, 48)
<i>24 months</i>			
<i>n</i>	63	58	121
Mean (SD)	39.0 (10.40)	40.8 (10.34)	39.9 (10.37)
Median (min., max.)	43 (16, 48)	45 (1, 48)	44 (1, 48)
Neer three- and four-part fractures (<i>n</i> = 104)			
<i>6 months</i>			
<i>n</i>	44	52	96
Mean (SD)	34.0 (10.21)	31.1 (11.30)	32.4 (10.86)
Median (min., max.)	36.5 (11, 48)	32 (7, 48)	33 (7, 48)
<i>12 months</i>			
<i>n</i>	45	52	97
Mean (SD)	35.0 (11.03)	34.1 (10.80)	34.5 (10.86)
Median (min., max.)	37 (1, 48)	36 (10, 48)	37 (1, 48)
<i>24 months</i>			
<i>n</i>	43	51	94
Mean (SD)	37.2 (9.16)	35.6 (11.08)	36.3 (10.23)
Median (min., max.)	39 (16, 48)	39 (3, 48)	39 (3, 48)
Max., maximum; min., minimum.			

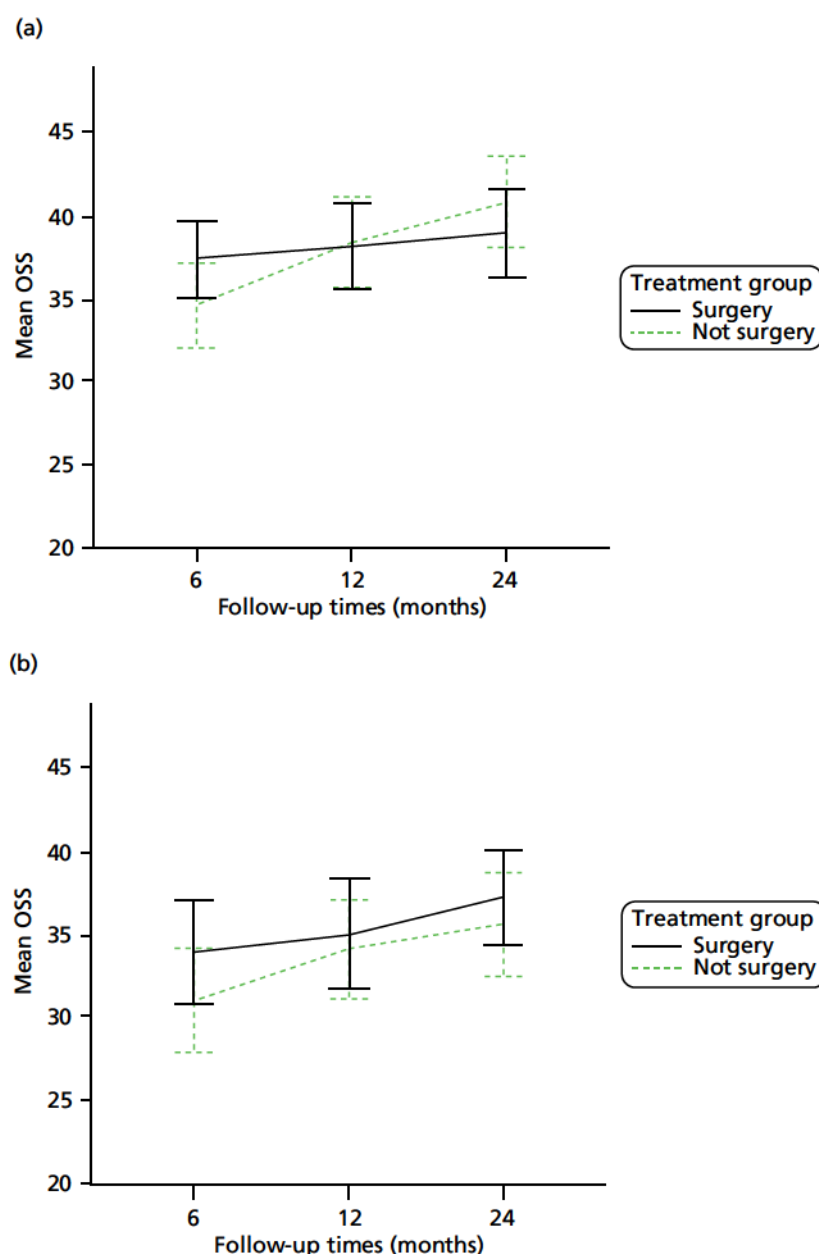


FIGURE 12 Mean OSS scores (with 95% CIs) over time by treatment group and Neer grouping. (a) Neer one- and two-part fractures and (b) Neer three- and four-part fractures.

A similar multilevel model to the primary analysis was applied but the Neer grouping term replaced the tuberosity involvement term and the model additionally included a treatment group \times Neer grouping interaction. Adjusted OSS means and group differences for this model are presented in *Table 33*. Similar to the primary analysis, there was no significant overall effect of treatment group over the 2-year follow-up period and no significant effect of treatment group at individual time points. A 2-log likelihood comparison with an adjusted primary analysis model (including Neer grouping instead of tuberosity involvement) was not significant [$\chi^2(1) = 0.05$, $p = 0.823$]; therefore, inclusion of the interaction did not significantly improve the model.

TABLE 33 Difference in mean OSS scores over time by treatment group: subgroup analysis 2 – Neer grouping (one- and two-part fractures/three- and four-part fractures)^a

Follow-up	Surgery (<i>n</i> = 125; <i>n</i> = 114 in analysis), mean (95% CI)	Not surgery (<i>n</i> = 125; <i>n</i> = 117 in analysis), mean (95% CI)	Difference (95% CI)	<i>p</i> -value
Overall	38.20 (36.50 to 39.82)	37.65 (36.00 to 39.23)	0.55 (−1.58 to 2.68)	0.613
6 months	36.93 (35.09 to 38.69)	34.86 (32.98 to 36.64)	2.08 (−0.29 to 4.45)	0.086
12 months	38.36 (36.58 to 40.06)	38.15 (36.42 to 39.80)	0.21 (−2.03 to 2.46)	0.853
24 months	39.26 (37.46 to 40.99)	39.77 (38.05 to 41.43)	−0.51 (−2.79 to 1.77)	0.661

a Multilevel model adjusted for treatment group, time (6, 12 and 24 months), treatment group × time interaction, baseline EQ-5D index, gender, age (< 65 years/≥ 65 years), Neer grouping (one- and two-part fractures/three- and four-part fractures) and treatment group × Neer grouping interaction.

Sensitivity analysis: smoking

A sensitivity analysis was undertaken for the primary outcome model to explore whether or not outcomes were robust when smoking status was taken into account. Smoking is associated with a number of complications and additionally was found to be imbalanced at baseline (surgery group: 19% smokers; not surgery group: 32% smokers).

Descriptive OSS statistics by treatment group and smoking status at 6, 12 and 24 months are provided in *Table 34* and illustrated in *Figure 13*.

The OSS scores improved for both smokers and non-smokers over the trial period. Because of the relatively smaller number of participants who were smokers, the variability of scores around the mean was greater for this group. Although scores at 6 and 12 months' follow-up were comparable between smokers and non-smokers, at 24 months the mean OSS score was around 2 score points higher for non-smokers. Again, surgical group scores were noticeably higher than those for the non-surgery group at 6 months, although CIs substantially overlapped.

The same multilevel model as for the primary analysis was applied but an additional smoking status was included as a covariate. Adjusted OSS means and group differences for this model are presented in *Table 35*. Similar to the primary analysis, there was no significant overall effect of treatment group over the 2-year follow-up period and no significant effect of treatment group at individual time points. Therefore, the results of the primary analysis are robust.

TABLE 34 Descriptive OSS statistics over time by treatment group and smoking status

Follow-up	Surgery (n = 125)	Not surgery (n = 125)	Total (n = 250)
Smoking status: smoker (n = 64)			
<i>6 months</i>			
<i>n</i>	22	37	59
Mean (SD)	34.7 (11.99)	31.7 (10.61)	32.8 (11.14)
Median (min., max.)	39 (11, 48)	33 (8, 48)	34 (8, 48)
<i>12 months</i>			
<i>n</i>	22	33	55
Mean (SD)	37.4 (11.68)	36.1 (10.76)	36.6 (11.05)
Median (min., max.)	42 (14, 48)	40 (13, 48)	42 (13, 48)
<i>24 months</i>			
<i>n</i>	20 (16.0)	33	53 (21.2)
Mean (SD)	36.8 (11.04)	37.0 (11.41)	36.9 (11.17)
Median (min., max.)	42 (16, 48)	41 (3, 48)	41 (3, 48)
Smoking status: non-smoker (n = 177)			
<i>6 months</i>			
<i>n</i>	86	78	164
Mean (SD)	36.3 (9.55)	33.4 (11.21)	34.9 (10.44)
Median (min., max.)	39 (14, 48)	36 (3, 48)	37 (3, 48)
<i>12 months</i>			
<i>n</i>	85	77	162
Mean (SD)	36.6 (10.77)	36.3 (11.01)	36.4 (10.85)
Median (min., max.)	40 (1, 48)	41 (4, 48)	41 (1, 48)
<i>24 months</i>			
<i>n</i>	82	72	154
Mean (SD)	38.7 (9.59)	38.7 (10.95)	38.7 (10.21)
Median (min., max.)	42 (16, 48)	42.5 (1, 48)	42 (1, 48)
Max., maximum; min., minimum.			

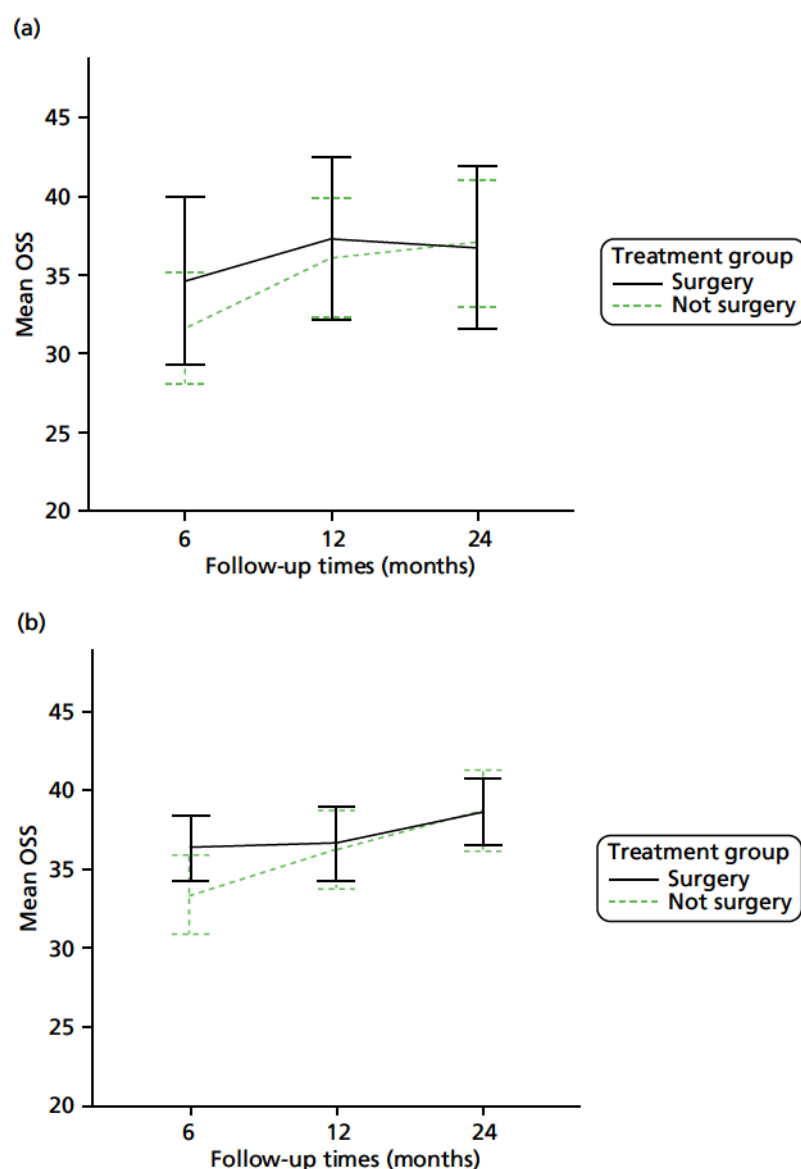


FIGURE 13 Mean OSS scores (with 95% CIs) over time by treatment group and smoking status. (a) Smoker and (b) non-smoker.

TABLE 35 Difference in mean OSS scores over time by treatment group: sensitivity analysis – smoking^a

Follow-up	Surgery (<i>n</i> = 125; <i>n</i> = 109 in analysis), mean (95% CI)	Not surgery (<i>n</i> = 125; <i>n</i> = 114 in analysis), mean (95% CI)	Difference (95% CI)	<i>p</i> -value
Overall	38.65 (36.74 to 40.48)	38.05 (36.24 to 39.79)	0.60 (−1.57 to 2.77)	0.588
6 months	37.40 (35.34 to 39.35)	35.31 (33.26 to 37.24)	2.09 (−0.32 to 4.50)	0.089
12 months	38.76 (36.76 to 40.66)	38.53 (36.65 to 40.33)	0.23 (−2.06 to 2.51)	0.847
24 months	39.77 (37.75 to 41.69)	40.16 (38.29 to 41.96)	−0.40 (−2.72 to 1.93)	0.739

^a Multilevel model adjusted for treatment group, time (6, 12 and 24 months), group × time interaction, smoking status (yes/no), baseline EQ-5D index, gender, age group (< 65 years/≥ 65 years) and tuberosity involvement at baseline (yes/no).

Patients' baseline treatment preferences

Treatment preference was recorded for all eligible patients. Surgeons recorded treatment preference for non-consenting patients, and consenting patients recorded whether or not they had any treatment preference at baseline, before randomisation.

Patient preferences at baseline are shown in *Table 36*. The majority of non-consenting patients (72%) expressed a preference for non-surgical treatment, whereas this applied to only 24% of consenters. Of the consenters, just under half (46%) had no treatment preference and the remainder were split slightly in favour of surgical treatment (29% vs. 24% for non-surgical treatment). Patients in the surgery group showed a greater preference for surgical treatment (33% vs. 25% in the non-surgery group), whereas patients in the non-surgery group were more likely to have no preference (50% vs. 42% in the surgery group).

Descriptive OSS statistics by treatment group and patient preference at 6, 12 and 24 months are detailed in *Table 37* and illustrated in *Figure 14*. Although OSS scores improved for all groups over time, patients who expressed a preference for surgery generally reported the lowest OSS scores followed by patients who had no preference; patients who expressed a preference for no surgery had the highest OSS scores. The difference between treatment groups in favour of surgery at 6 months' follow-up was most pronounced for patients who had no preference for either treatment.

Patients who preferred surgery tended to be older (by 3 years on average) and presented earlier (by 2 days) with more complex fractures (7% more fractures involving one or both tuberosities) than patients who preferred not to have surgery; this gives an indication why these patients had poorer OSS outcomes.

The same multilevel model as for the primary analysis was applied but additionally patient preference at baseline and its interaction with allocated treatment group were included as covariates. Adjusted OSS means and group differences are presented in *Table 38*. Again, there was no significant overall effect of treatment group over the 2-year follow-up period or at individual time points.

Of all the analyses adding further terms to the primary model (age, fracture type, smoking status, patient preference), this model reduced the magnitude of the treatment effect the most (from an overall group difference of 0.75 score points to 0.50 score points), reflecting the additional variability in OSS scores explained by patient preference, as seen in *Figure 14*. However, the interaction between patient preference and allocated treatment was not statistically significant ($F = 0.29$, $p = 0.751$), even against a relaxed significance level of 0.1.

TABLE 36 Treatment preference of eligible patients at baseline

Preference	Non-consenters ($n = 313$), n (%)	Consenters, n (%)		
		Surgery ($n = 125$)	Not surgery ($n = 125$)	Total ($n = 250$)
Surgery	55 (17.6)	41 (32.8)	31 (24.8)	72 (28.8)
Not surgery	226 (72.2)	32 (25.6)	28 (22.4)	60 (24.0)
No preference	23 (7.3)	52 (41.6)	63 (50.4)	115 (46.0)
Missing data	9 (2.9)	0 (0.0)	3 (2.4)	3 (1.2)

TABLE 37 Descriptive OSS statistics over time by treatment group and patient preference

Follow-up	Surgery (n = 125)	Not surgery (n = 125)	Total (n = 250)
Preference surgery (n = 72)			
<i>6 months</i>			
<i>n</i>	39	31	70
Mean (SD)	32.9 (11.38)	31.7 (12.80)	32.4 (11.96)
Median (min., max.)	33 (11, 48)	32 (3, 48)	33 (3, 48)
<i>12 months</i>			
<i>n</i>	38	30	68
Mean (SD)	34.1 (11.50)	34.4 (12.10)	34.2 (11.68)
Median (min., max.)	36 (1, 48)	37.5 (4, 48)	37 (1, 48)
<i>24 months</i>			
<i>n</i>	37	27	64
Mean (SD)	36.2 (11.61)	36.0 (11.92)	36.1 (11.65)
Median (min., max.)	41 (16, 48)	39 (1, 48)	39.5 (1, 48)
Preference not surgery (n = 60)			
<i>6 months</i>			
<i>n</i>	28	27	55
Mean (SD)	39.6 (8.09)	37.0 (7.29)	38.3 (7.75)
Median (min., max.)	42.5 (20, 48)	39 (20, 48)	39 (20, 48)
<i>12 months</i>			
<i>n</i>	27	27	54
Mean (SD)	39.6 (8.67)	39.7 (7.14)	39.6 (7.86)
Median (min., max.)	41 (19, 48)	42 (23, 48)	41.5 (19, 48)
<i>24 months</i>			
<i>n</i>	25	27	52
Mean (SD)	41.5 (7.07)	42.2 (6.16)	41.9 (6.55)
Median (min., max.)	44 (23, 48)	44 (25, 48)	44 (213, 48)
No preference (n = 115)			
<i>6 months</i>			
<i>n</i>	46	58	104
Mean (SD)	36.6 (9.11)	31.9 (11.28)	34.0 (10.60)
Median (min., max.)	39 (17, 48)	34.5 (7, 48)	37 (7, 48)
<i>12 months</i>			
<i>n</i>	46	54	100
Mean (SD)	37.7 (10.98)	36.2 (11.23)	36.9 (11.08)
Median (min., max.)	42 (5, 48)	42 (10, 48)	42 (5, 48)
<i>24 months</i>			
<i>n</i>	44	52	96
Mean (SD)	38.1 (9.44)	37.5 (12.18)	37.8 (10.96)
Median (min., max.)	41.5 (16, 48)	42 (3, 48)	42 (3, 48)
Max., maximum; min., minimum.			

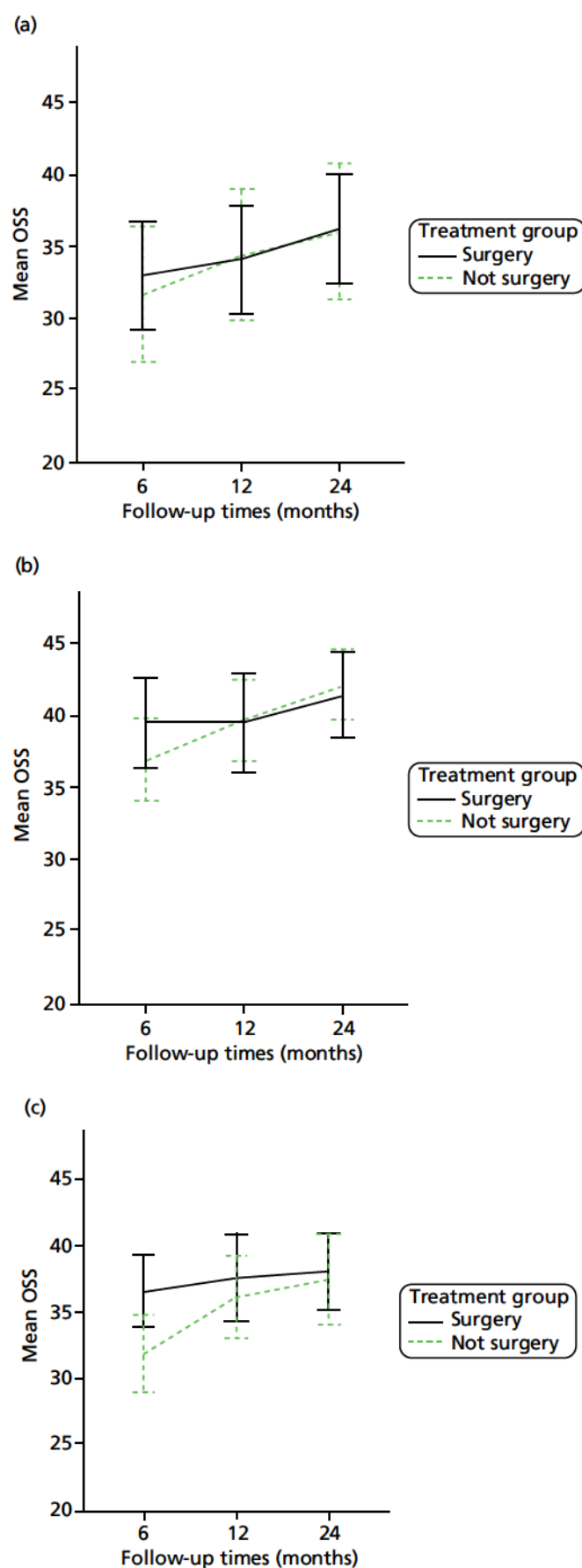


FIGURE 14 Mean OSS scores (with 95% CIs) over time by treatment group and patient preference. (a) Preference surgery; (b) preference not surgery; and (c) no preference.

TABLE 38 Difference in mean OSS scores over time by treatment group: including patient preferences at baseline^a

Follow-up	Surgery (<i>n</i> = 125; <i>n</i> = 114 in analysis), mean (95% CI)	Not surgery (<i>n</i> = 125; <i>n</i> = 115 in analysis), mean (95% CI)	Difference (95% CI)	<i>p</i> -value
Overall	39.14 (37.38 to 40.83)	38.65 (36.85 to 40.37)	0.50 (−1.64 to 2.63)	0.649
6 months	37.91 (36.01 to 39.72)	35.91 (33.89 to 37.83)	2.00 (−0.37 to 4.36)	0.098
12 months	39.30 (37.46 to 41.06)	39.22 (37.36 to 40.99)	0.86 (−2.16 to 2.33)	0.940
24 months	40.19 (38.32 to 41.98)	40.66 (38.80 to 42.44)	−0.47 (−2.76 to 1.82)	0.687

a Multilevel model adjusted for treatment group, time (6, 12 and 24 months), group × time interaction, patient preference, group × patient preference interaction, baseline EQ-5D index, gender, age group (< 65 years/≥ 65 years) and tuberosity involvement at baseline (yes/no).

Exploration of centre variability

In our analysis plan we revised our intention to explore the potential for clustering, from individual surgeons to the level of the centre, in essence quantifying the performance of the surgical team(s) rather than that of an individual surgeon within the centre. This was because of the low numbers of patients treated by an individual surgeon. Among the 33 centres that screened patients, around 60% (*n* = 148) of the 250 participants were recruited by eight sites (12–43 patients each, 18.5 patients on average) and around 40% (*n* = 102) were recruited by the remaining 25 centres (0–9 patients each, 4.1 patients on average).

Mean OSS scores and 95% CIs for randomised patients in each centre at each follow-up time point are illustrated in *Figure 15*. These ranged from 22 to 48 score points, but no centre appeared to be substantially outlying. When CIs were either exceptionally wide or did not include the overall mean, this was predominantly because of the small number of patients in those centres. Extreme CI limits are not displayed in full here. When no CIs are displayed, the sample size was equal to 1, except in one case in which there were two patients with identical scores.

The distribution of scores is further examined using box plots of OSS scores (*Figure 16*). These show a wide spread of scores within each centre, again affected by small sample sizes in some centres, and highlight a small number of outliers.

Summary of the Oxford Shoulder Score multilevel modelling results

Table 39 summarises the OSS estimates for each treatment group as well as group differences following multilevel analysis. These are generated using a mixed-effect model with a random slope of time within patients. The outcome is the squared OSS (estimates are back transformed in the table). Fixed effects for each model are given in the footnotes.

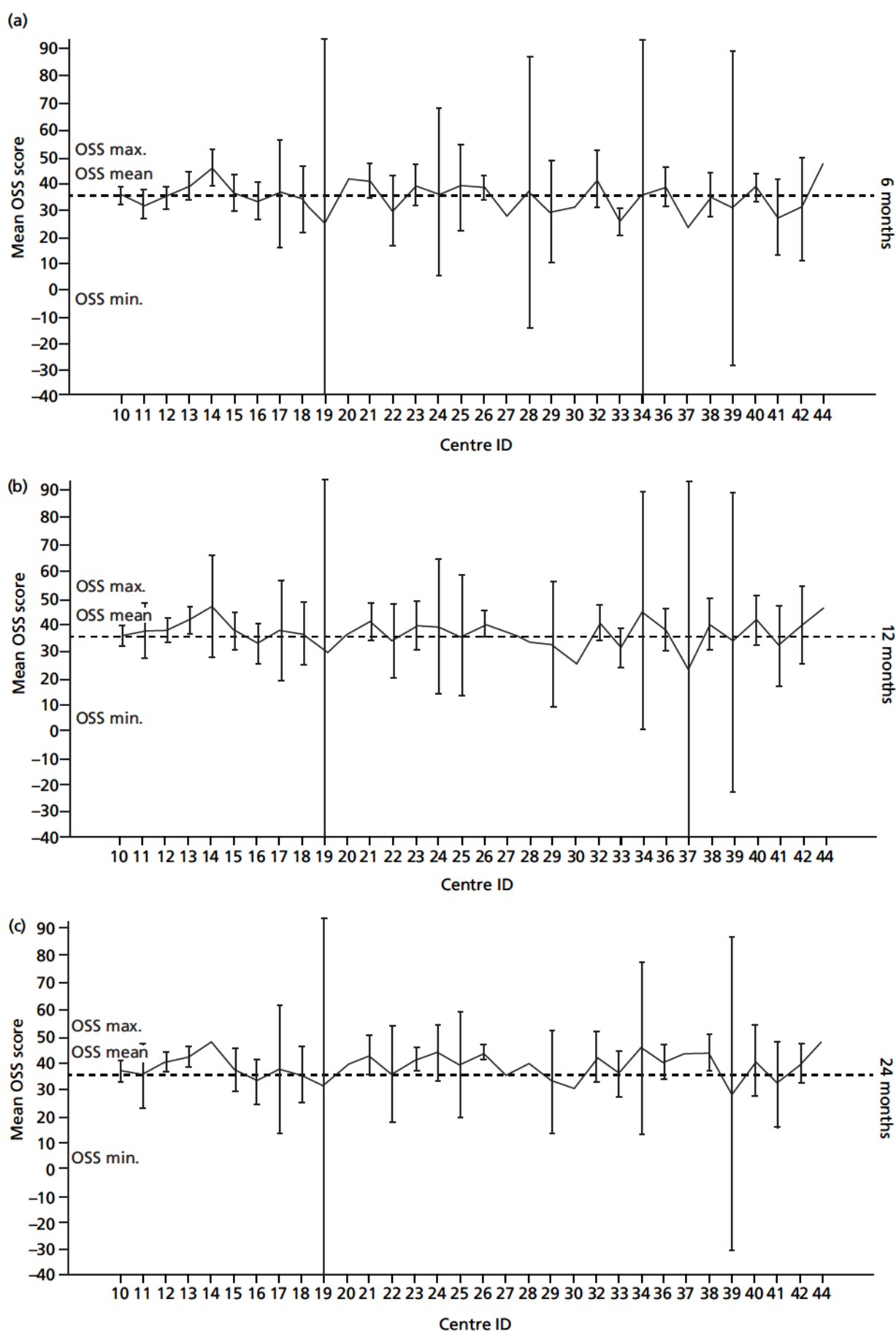


FIGURE 15 Mean OSS scores (with 95% CIs) over time by recruitment site. Max., maximum; min., minimum.

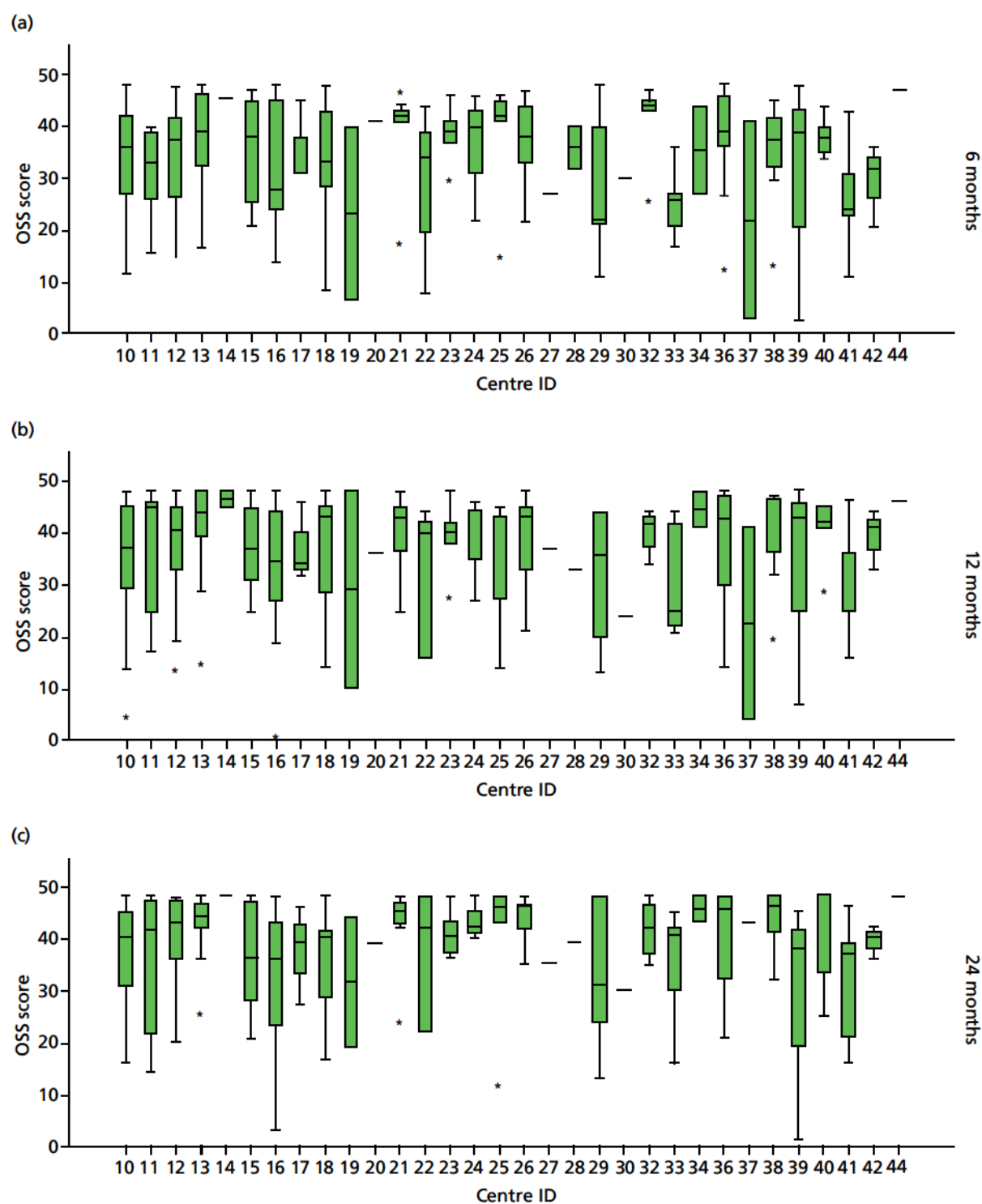


FIGURE 16 Distribution of OSS scores (box plots) by recruitment site.

TABLE 39 Difference in mean OSS scores over time by treatment group

Follow-up	Surgery (<i>n</i> = 125), ^a mean (95% CI)	Not surgery (<i>n</i> = 125), ^a mean (95% CI)	Difference (95% CI)	<i>p</i> -value
Primary analysis ^b				
Overall	39.07 (37.30 to 40.76)	38.32 (36.57 to 39.99)	0.75 (−1.33 to 2.84)	0.479
6 months	37.84 (35.93 to 39.65)	35.59 (33.62 to 37.45)	2.25 (−0.07 to 4.57)	0.058
12 months	39.23 (37.38 to 40.99)	38.80 (36.99 to 40.53)	0.42 (−1.78 to 2.63)	0.706
24 months	40.11 (38.24 to 41.90)	40.40 (38.59 to 42.13)	−0.29 (−2.53 to 1.95)	0.800
Subgroup analysis 1: age (<65 years/≥ 65 years) ^c				
Overall	38.98 (37.21 to 40.68)	38.40 (36.66 to 40.07)	0.58 (−1.52 to 2.68)	0.590
6 months	37.74 (35.83 to 39.56)	35.69 (33.72 to 37.55)	2.06 (−0.28 to 4.40)	0.085
12 months	39.14 (37.29 to 40.91)	38.89 (37.08 to 40.62)	0.26 (−1.96 to 2.47)	0.822
24 months	40.03 (38.15 to 41.82)	40.48 (38.68 to 42.21)	−0.46 (−2.71 to 1.79)	0.690
Subgroup analysis 2: tuberosity involvement (yes/no) ^d				
Overall	39.05 (37.10 to 40.90)	38.33 (36.48 to 40.10)	0.71 (−1.79 to 3.22)	0.577
6 months	37.81 (35.73 to 39.79)	35.61 (33.53 to 37.57)	2.20 (−0.54 to 4.94)	0.115
12 months	39.20 (37.18 to 41.12)	38.82 (36.91 to 40.64)	0.38 (−2.22 to 2.98)	0.772
24 months	40.09 (38.05 to 42.02)	40.42 (38.52 to 42.23)	−0.33 (−2.94 to 2.28)	0.804
Subgroup analysis 2 (sensitivity analysis): Neer grouping (one- and two-part fractures/three- and four-part fractures) ^e				
Overall	38.20 (36.50 to 39.82)	37.65 (36.00 to 39.23)	0.55 (−1.58 to 2.68)	0.613
6 months	36.93 (35.09 to 38.69)	34.86 (32.98 to 36.64)	2.08 (−0.29 to 4.45)	0.086
12 months	38.36 (36.58 to 40.06)	38.15 (36.42 to 39.80)	0.21 (−2.03 to 2.46)	0.853
24 months	39.26 (37.46 to 40.99)	39.77 (38.05 to 41.43)	−0.51 (−2.79 to 1.77)	0.661
Sensitivity analysis: smoking status (yes/no) ^f				
Overall	38.65 (36.74 to 40.48)	38.05 (36.24 to 39.79)	0.60 (−1.57 to 2.77)	0.588
6 months	37.40 (35.34 to 39.35)	35.31 (33.26 to 37.24)	2.09 (−0.32 to 4.50)	0.089
12 months	38.76 (36.76 to 40.66)	38.53 (36.65 to 40.33)	0.23 (−2.06 to 2.51)	0.847
24 months	39.77 (37.75 to 41.69)	40.16 (38.29 to 41.96)	−0.40 (−2.72 to 1.93)	0.739
				continued

TABLE 39 Difference in mean OSS scores over time by treatment group (*continued*)

Follow-up	Surgery (<i>n</i> = 125), ^a mean (95% CI)	Not surgery (<i>n</i> = 125), ^a mean (95% CI)	Difference (95% CI)	<i>p</i> -value
Patient preference analysis (surgery/not surgery/no preference) ^g				
Overall	39.14 (37.38 to 40.83)	38.65 (36.85 to 40.37)	0.50 (−1.64 to 2.63)	0.649
6 months	37.91 (36.01 to 39.72)	35.91 (33.89 to 37.83)	2.00 (−0.37 to 4.36)	0.098
12 months	39.30 (37.46 to 41.06)	39.22 (37.36 to 40.99)	0.86 (−2.16 to 2.33)	0.940
24 months	40.19 (38.32 to 41.98)	40.66 (38.80 to 42.44)	−0.47 (−2.76 to 1.82)	0.687

a Numbers given are those randomised. The numbers of patients included in each analysis are stated in the footnotes. These are the numbers of patients with valid OSS data for at least one follow-up time point and complete baseline covariates.

b Fixed effects: group, time (6, 12 and 24 months), group × time interaction, baseline EQ-5D index, gender, age (< 65 years/≥ 65 years), tuberosity involvement at baseline (yes/no). Included in the analysis: 114 surgery group, 117 not surgery group.

c Fixed effects: group, time (6, 12 and 24 months), group × time interaction, baseline EQ-5D index, gender, age group (< 65 years/≥ 65 years), group × age group interaction, tuberosity involvement at baseline (yes/no). Included in the analysis: 114 surgery group, 117 not surgery group.

d Fixed effects: group, time (6, 12 and 24 months), group × time interaction, baseline EQ-5D index, gender, age group (< 65 years/≥ 65 years), tuberosity involvement at baseline (yes/no), group × tuberosity involvement interaction. Included in the analysis: 114 surgery group, 117 not surgery group.

e Fixed effects: group, time (6, 12 and 24 months), group × time interaction, baseline EQ-5D index, gender, age group (< 65 years/≥ 65 years), Neer grouping (one- and two-part fractures/three- and four-part fractures), group × Neer grouping interaction. Included in the analysis: 114 surgery group, 117 not surgery group.

f Additional fixed effects: group, time (6, 12 and 24 months), group × time interaction, baseline EQ-5D index, gender, age group (< 65 years/≥ 65 years), smoking status (yes/no), tuberosity involvement at baseline (yes/no). Included in the analysis: 109 surgery group, 114 not surgery group.

g Additional fixed effects: group, time (6, 12 and 24 months), group × time interaction, baseline EQ-5D index, gender, age group (< 65 years/≥ 65 years), patient preference, group × patient preference interaction, tuberosity involvement at baseline (yes/no). Included in the analysis: 114 surgery group, 115 not surgery group.

Chapter 6 Results: analyses of secondary and other outcomes and supplementary analyses

12-item Short Form health survey

The 12 SF-12 questions were summarised into eight domain scores, each based on one or two questionnaire items. If at least one question had been completed for a domain, any missing items were replaced with the available score for that domain. Weighted estimates of the eight SF-12 domains were then used to calculate the PCS score and MCS score, each ranging from 0 to 100 score points, with higher scores representing a better health status. All domain scores needed to be available for the calculation of the PCS and MCS scores. Scores were standardised against the general US population (mean 50, SD 10).⁶⁹ Using standard effect sizes, differences of 2, 5 and 8 score points would correspond to small, medium and large differences respectively.

Physical component summary score

Descriptive SF-12 PCS score statistics at 6, 12 and 24 months are shown in *Table 40* and illustrated in *Figure 17*. The distribution of scores is presented in *Figure 18*.

Overall, physical functioning remained largely unchanged over the trial period. The distribution of scores was approximately normal at all time points. Mean PCS scores were slightly higher for the surgery group at all time points, with group differences being largest at 6 months. There was substantial overlap in the CIs between groups.

A similar multilevel model to that in the OSS analysis was applied, using SF-12 PCS score as the outcome (no transformation was found to be necessary) and including the same predictors. Adjusted SF-12 PCS score means and group differences for this model are presented in *Table 41*. The results mirrored those of the OSS analysis, with slightly better functioning in the surgery group by an average of between 1.3 and 2.6 score points over the 2-year follow-up period. None of the differences between the two groups, overall or at individual time points, was statistically significant. Similar to the OSS outcome, the largest group difference was observed at 6 months' follow-up.

Mental component summary score

Descriptive SF-12 MCS score statistics at 6, 12 and 24 months are shown in *Table 42* and illustrated in *Figure 19*. The distribution of scores is presented in *Figure 20*.

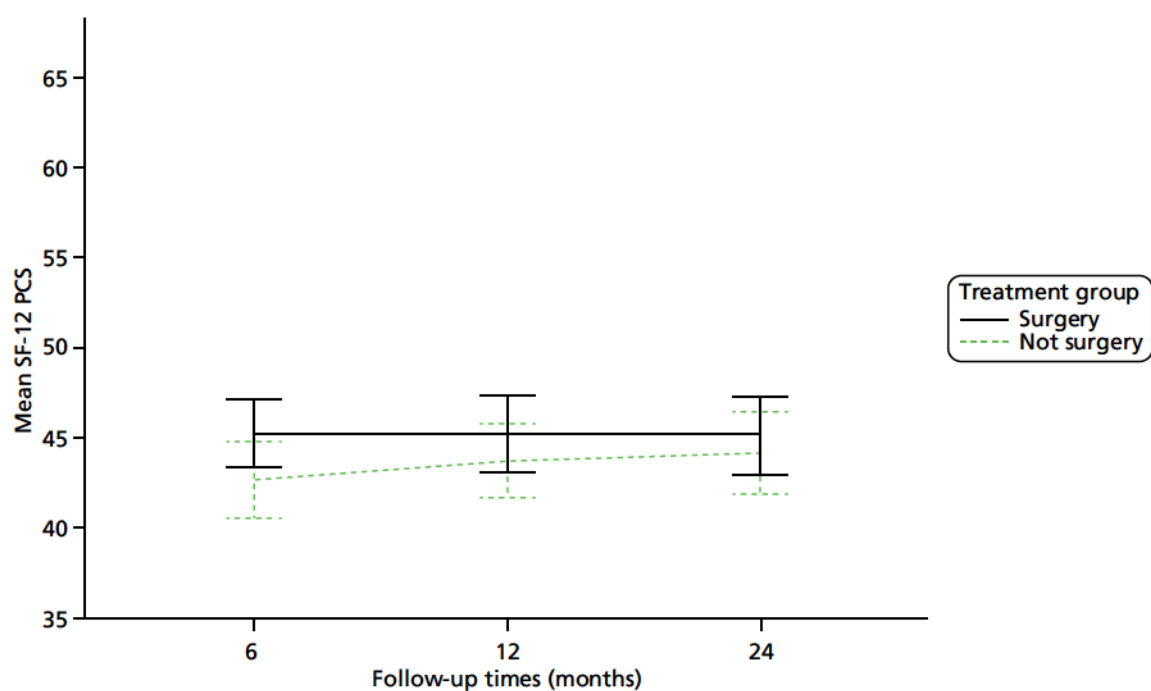
Overall, mental functioning improved slightly over the trial period. The distribution of scores was slightly left skewed at 6 and 12 months, which became more pronounced at 24 months. Unlike physical functioning, as assessed by the OSS and PCS score, the mean MCS score was slightly higher for the non-surgery group at all time points, with the group difference being greatest at 12 months. There was substantial overlap of the 95% CIs of the two groups.

A similar multilevel model to the OSS analysis was applied, using SF-12 MCS score as the outcome (no transformation was found to be necessary) and including the same predictors. Adjusted SF-12 MCS score means and group differences for this model are presented in *Table 43*. There were better scores in the non-surgery arm by an average of between 0.5 and 2.0 score points over the 2-year follow-up period. This difference was not statistically significant over the 2-year follow-up period or at individual time points.

TABLE 40 Descriptive SF-12 PCS score statistics over time by treatment group

Follow-up	Surgery (n = 125)	Not surgery (n = 125)	Total (N = 250)
6 months			
n (%)	106 (84.8)	110 (88.0)	216 (86.4)
Mean (SD)	45.3 (10.01)	42.7 (11.25)	43.9 (10.71)
Median (min., max.)	46.7 (15.4, 64.1)	43.6 (11.9, 62.4)	45.0 (11.9, 64.1)
12 months			
n (%)	108 (86.4)	110 (88.0)	218 (87.2)
Mean (SD)	45.2 (10.98)	43.7 (10.98)	44.5 (10.93)
Median (min., max.)	45.2 (14.9, 63.7)	44.1 (17.6, 61.3)	44.7 (14.9, 63.7)
24 months			
n (%)	105 (84.0)	105 (84.0)	210 (84.0)
Mean (SD)	45.2 (11.30)	44.1 (11.58)	44.6 (11.43)
Median (min., max.)	46.1 (14.9, 65.7)	45.2 (14.3, 60.7)	46.0 (14.3, 65.7)

Max., maximum; min., minimum.

**FIGURE 17** Mean SF-12 PCS scores (with 95% CIs) over time by treatment group.

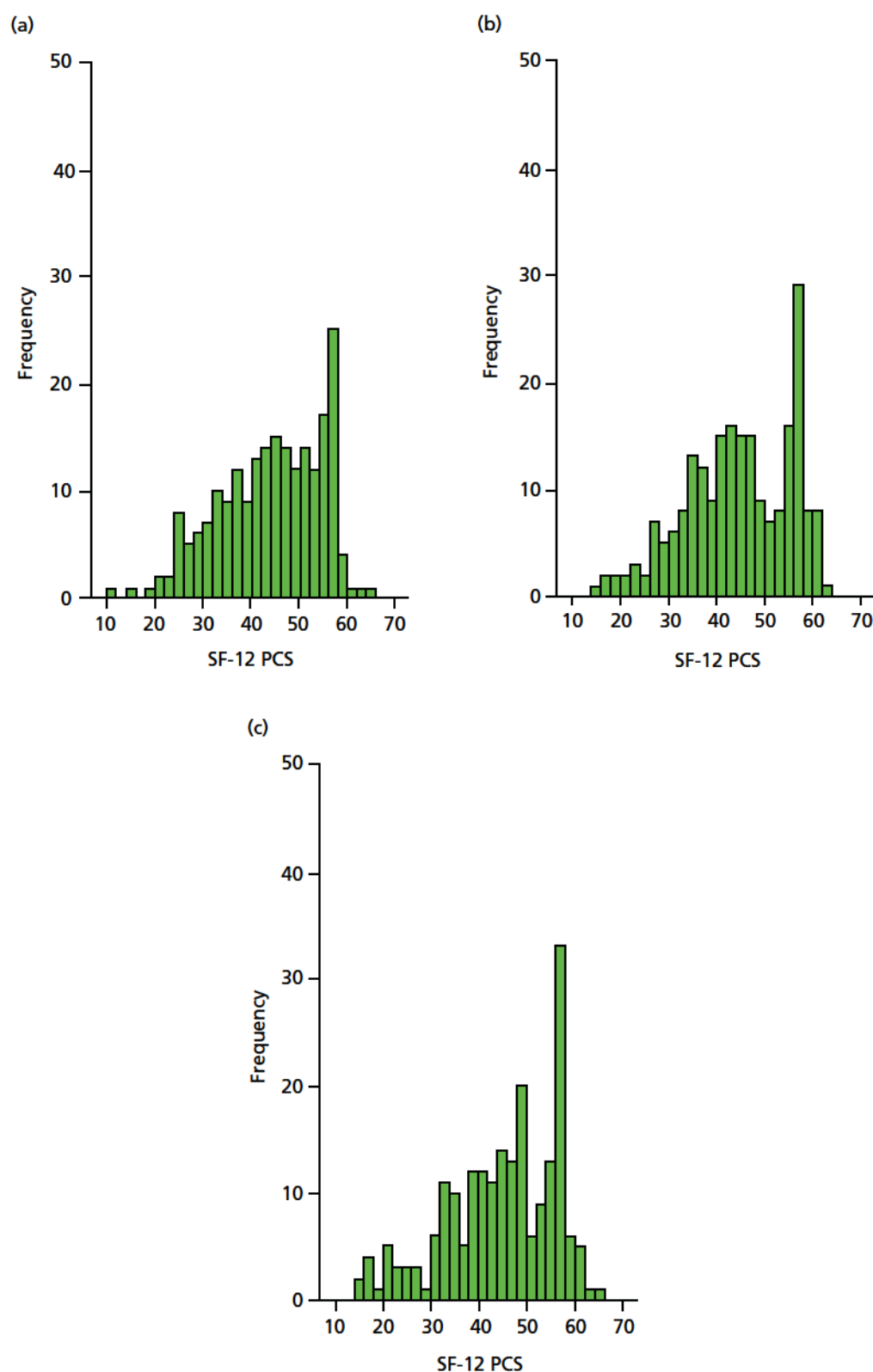


FIGURE 18 Distribution of SF-12 PCS scores over time. (a) 6 months; (b) 12 months; and (c) 24 months.

TABLE 41 Difference in mean SF-12 PCS scores over time by treatment group^a

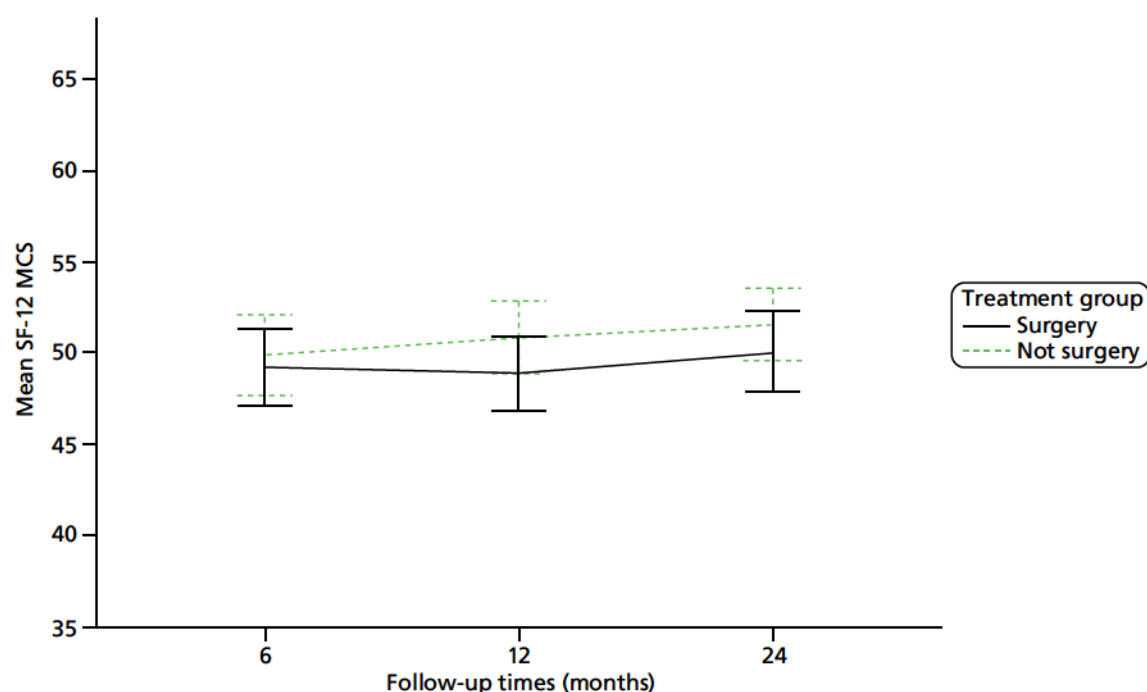
Follow-up	Surgery (<i>n</i> = 125; <i>n</i> = 111 in analysis), mean (95% CI)	Not surgery (<i>n</i> = 125; <i>n</i> = 115 in analysis), mean (95% CI)	Difference (95% CI)	<i>p</i> -value
Overall	45.64 (43.44 to 47.84)	43.87 (41.75 to 45.99)	1.77 (−0.84 to 4.39)	0.184
6 months	45.73 (43.44 to 48.02)	43.18 (40.97 to 45.39)	2.55 (−0.21 to 5.32)	0.070
12 months	45.51 (43.22 to 47.80)	44.22 (42.01 to 46.43)	1.29 (−1.48 to 4.06)	0.361
24 months	45.68 (43.28 to 48.08)	44.20 (41.87 to 46.54)	1.48 (−1.48 to 4.43)	0.327

a Multilevel model adjusted for treatment group, time (6, 12 and 24 months), group × time interaction, baseline EQ-5D index, gender, age group (< 65 years/≥ 65 years) and tuberosity involvement at baseline (yes/no).

TABLE 42 Descriptive SF-12 MCS score statistics over time by treatment group

Follow-up	Surgery (<i>n</i> = 125)	Not surgery (<i>n</i> = 125)	Total (<i>N</i> = 250)
6 months			
<i>n</i> (%)	106 (84.8)	110 (88.0)	216 (86.4)
Mean (SD)	49.2 (10.84)	49.8 (11.46)	49.5 (11.14)
Median (min., max.)	52.2 (20.3, 63.7)	52.1 (17.6, 68.4)	52.2 (17.6, 68.4)
12 months			
<i>n</i> (%)	108 (86.4)	110 (88.0)	218 (87.2)
Mean (SD)	48.8 (10.51)	50.8 (10.67)	49.8 (10.61)
Median (min., max.)	51.2 (17.8, 66.2)	54.0 (15.8, 67.4)	52.6 (15.8, 67.4)
24 months			
<i>n</i> (%)	105 (84.0)	105 (84.0)	210 (84.0)
Mean (SD)	50.1 (11.64)	51.5 (9.96)	50.8 (10.83)
Median (min., max.)	55.2 (8.8, 64.4)	54.2 (21.2, 66.1)	54.3 (8.8, 66.1)

Max., maximum; min., minimum.

**FIGURE 19** Mean SF-12 MCS scores (with 95% CIs) over time by treatment group.

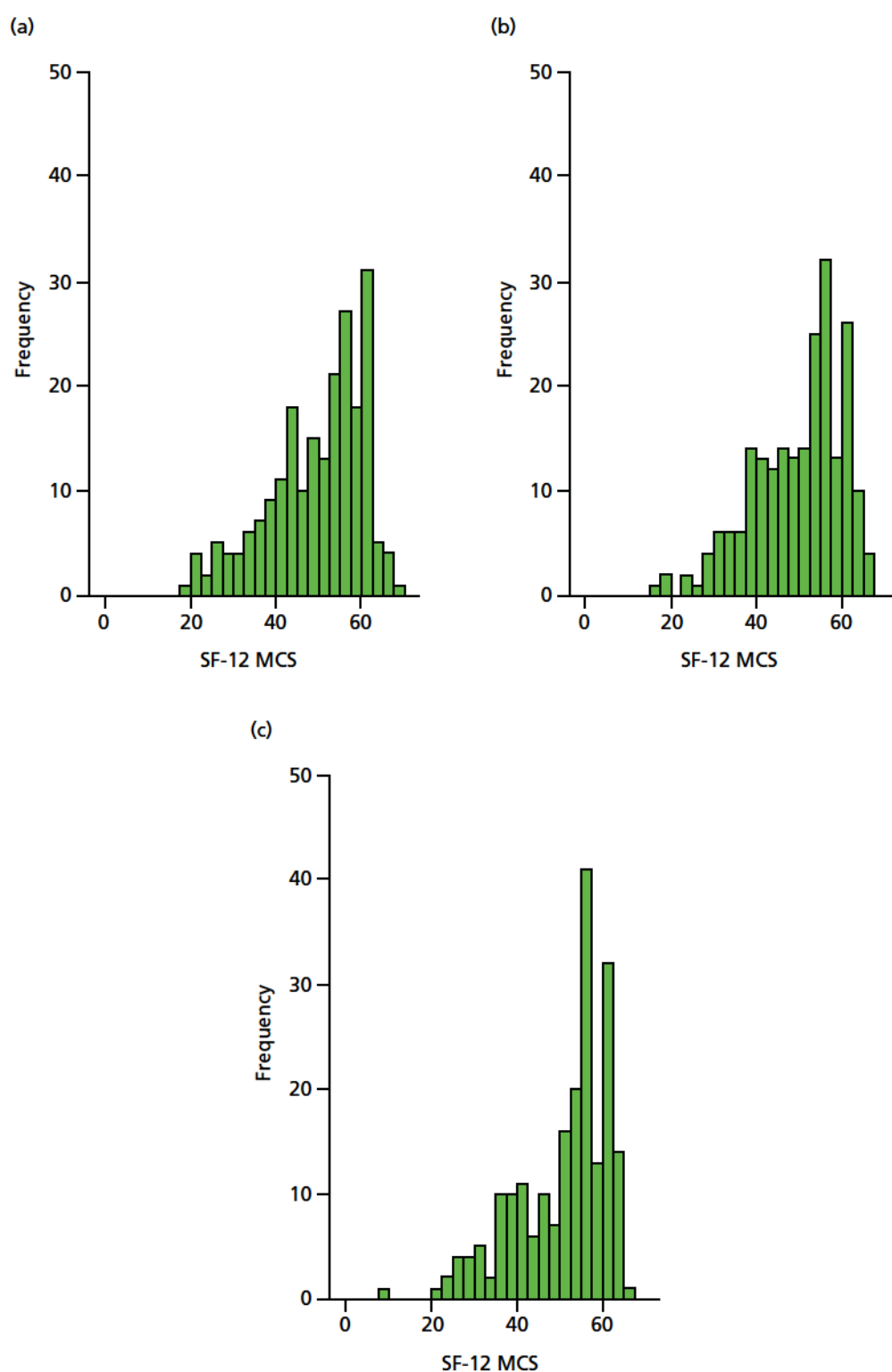


FIGURE 20 Distribution of SF-12 MCS scores over time. (a) 6 months; (b) 12 months; and (c) 24 months.

TABLE 43 Difference in mean SF-12 MCS scores over time by treatment group^a

Follow-up	Surgery (<i>n</i> = 125; <i>n</i> = 111 in analysis), mean (95% CI)	Not surgery (<i>n</i> = 125; <i>n</i> = 115 in analysis), mean (95% CI)	Difference (95% CI)	<i>p</i> -value
Overall	48.66 (46.55 to 50.77)	49.96 (47.92 to 52.00)	-1.28 (-3.80 to 1.23)	0.317
6 months	48.43 (46.07 to 50.80)	48.95 (46.66 to 51.24)	-0.52 (-3.44 to 2.41)	0.729
12 months	48.24 (45.96 to 50.53)	50.20 (47.98 to 52.41)	-1.95 (-4.76 to 0.85)	0.172
24 months	49.30 (46.97 to 51.64)	50.69 (48.40 to 52.97)	-1.38 (-4.27 to 1.51)	0.349

a Multilevel model adjusted for treatment group, time (6, 12 and 24 months), group × time interaction, baseline EQ-5D index, gender, age group (< 65 years/≥ 65 years) and tuberosity involvement at baseline (yes/no).

Surgical and shoulder fracture-related complications

Surgical and other shoulder fracture-related complications were recorded at the end of patients' orthopaedic inpatient episodes when applicable and at 1- and 2 years' follow-up. Clinic staff could select from a list of predefined complication categories or indicate 'other'. For each, descriptions could be added as free text and resulting treatments could be detailed in a further section on each form.

As described in *Chapter 3*, to ensure the validity and interpretability of the entered complications, two expert raters assessed each reported complication in the context of the information provided elsewhere on the same form or previous/subsequent forms as well as the other trial data available for that patient and agreed on any changes. Using clinical judgement, some records were recategorised and entries under 'other' were grouped when possible. Duplicate and irrelevant information was removed and additional complications were appended to the data if these were clearly reported elsewhere, for example when included in the description of further treatment (see *Appendix 36*).

The final data set consisted of 59 surgical and shoulder fracture-related complications reported for 53 patients (*Table 44*).

Apart from haematoma formation at the surgical site, all predefined complications were reported at least once. In the surgery group, a total of 30 patients (24%) experienced at least one complication over the 2-year follow-up period. These included six patients with two complications, which were concurrent for three of these patients. The majority of complications were reported at the 1-year follow-up. The most common complications in the surgery group were metalwork problems (*n* = 10) and post-traumatic stiffness (*n* = 6). It is noteworthy that three of the four reported malunions occurred for crossover patients who were allocated to surgery but who did not receive the intervention. In the non-surgery group, a total of 23 patients (18%) experienced one complication each over the trial period. All but four of these were reported at the 1-year follow-up. The most common complications in this group were malunion (*n* = 5), nonunion (*n* = 5) and post-traumatic stiffness (*n* = 5). Although marginally more patients in the surgery group experienced these complications, a chi-squared test revealed that there was no statistically significant difference between the treatment groups [$\chi^2(1) = 1.17$, *p* = 0.279].

To ascertain whether or not certain attributes may predispose patients to experience a surgical or shoulder-related complication, selected baseline characteristics for patients who suffered at least one complication (*n* = 53) and patients for whom no complications were reported (*n* = 197) were compared (*Table 45*). No marked between-group differences were evident. [Although patients with complications appeared to be marginally younger (49% were aged < 65 years compared with 42% of patients without complications), the mean ages of the treatment groups were similar.]

TABLE 44 Numbers of surgical and shoulder fracture-related complications by treatment group

Complication	Surgery (<i>n</i> = 125)			Not surgery (<i>n</i> = 125)			Total (<i>N</i> = 250)					
	Inpatient ^a	Year 1	Year 2	Total	Inpatient ^a	Year 1	Year 2	Total	Inpatient ^a	Year 1	Year 2	Total
Surgical site infection, <i>n</i>	1	1		2				0	1	1		2
Haematoma formation at surgical site, <i>n</i>				0				0				0
Nerve injury, <i>n</i>	1	1		2				0	1	1		2
AVN, <i>n</i>		2	2	4		1		1		3	2	5
Implant-related complication/failure, <i>n</i>		2		2				0		2		2
Dislocation/Instability, <i>n</i>				0		1 ^{xb}		1		1		1
Metalwork problems, <i>n</i>	2	5	3	10				0	2	5	3	10
Malunion, <i>n</i>		2 ^{x,xb}	2 ^{xb}	4		3	2	5		5	4	9
Nonunion, <i>n</i>				0		4	1	5		4	1	5
Other, <i>n</i>												
Post-traumatic stiffness		5	1	6		5		5		10	1	11
Rotator cuff tear		2	1	3		1		1		3	1	4
Complex regional pain syndrome		1		1				0		1		1
(Severe) pain			1	1		1		1		1	1	2
Impingement				0			1	1			1	1
Unclear			1	1		3		3		3	1	4
Total complications, <i>n</i>	4	21	11	36	0	19	4	23	4	40	15	59
Total patients, <i>n</i>	4	18	10	30 ^c	0	19	4	23	4	37	14	53 ^c
Total patients, %	3.2	14.4	8.0	24.0	0.0	15.2	3.2	18.4	1.6	14.8	5.6	21.2
Mean complications per patient, <i>n</i>	0.03	0.17	0.09	0.29	0.00	0.15	0.03	0.18	0.02	0.16	0.06	0.24

^a At the end of the orthopaedic inpatient episode.

^b Each x denotes that a crossover patient is included in the count, that is, the patient did not receive the allocated treatment.

^c Six patients in the surgical group experienced more than one complication (two in each case): surgical site infection + metalwork problems in year 1; metalwork problems during the inpatient stay + rotator cuff tear in year 2; AVN in year 1 + metalwork problems in year 2; concurrent metalwork problems + rotator cuff tear in year 1; concurrent malunion + post-traumatic stiffness in year 2; concurrent post-traumatic stiffness + rotator cuff tear in year 1.

TABLE 45 Baseline characteristics by presence and absence of surgical and shoulder fracture-related complications

Characteristic	Patients with at least one complication (<i>n</i> = 53)	Patients without complications (<i>n</i> = 197)	Total (<i>N</i> = 250)
Gender, <i>n</i> (%)			
Male	12 (22.6)	46 (23.4)	58 (23.2)
Female	41 (77.4)	151 (76.6)	192 (76.8)
Age (years)			
<i>n</i>	53	197	250
Mean (SD)	65.3 (13.33)	66.2 (11.56)	66.0 (11.94)
Median (min., max.)	65.2 (24.6, 89.0)	67.4 (31.3, 92.0)	66.9 (24.6, 92.0)
Age (group), <i>n</i> (%)			
< 65 years	26 (49.1)	82 (41.6)	108 (43.2)
≥ 65 years	27 (50.9)	115 (58.4)	142 (56.8)
Diabetes, <i>n</i> (%)			
Yes	7 (13.2)	24 (12.2)	31 (12.4)
No	46 (86.8)	171 (86.8)	217 (86.8)
Missing	0 (0.0)	2 (1.0)	2 (0.8)
Smoking status, <i>n</i> (%)			
Yes	15 (28.3)	49 (24.9)	64 (25.6)
No	37 (69.8)	140 (71.1)	177 (70.8)
Missing	1 (1.9)	8 (4.1)	9 (3.6)
Steroid use, <i>n</i> (%)			
Yes	3 (5.7)	10 (5.1)	13 (5.2)
No	49 (92.5)	185 (93.9)	234 (93.6)
Missing	1 (1.9)	2 (1.0)	3 (1.2)
Health status (EQ-5D index)			
<i>n</i>	51	193	244
Mean (SD)	0.40 (0.38)	0.41 (0.37)	0.40 (0.37)
Median (min., max.)	0.59 (−0.36, 1)	0.33 (−0.35, 1)	0.33 (−0.36, 1)
Tuberosity involvement, <i>n</i> (%)			
Yes	40 (75.5)	153 (77.7)	193 (77.2)
No	13 (24.5)	44 (22.3)	57 (22.8)
Tuberosity involvement, <i>n</i> (%)			
Tuberosity not involved or missing	13 (24.5)	44 (22.3)	57 (22.8)
Greater tuberosity	22 (41.5)	97 (49.2)	119 (47.6)
Lesser tuberosity	3 (5.7)	7 (3.6)	10 (4.0)
Greater and lesser tuberosity	15 (28.3)	49 (24.9)	64 (25.6)

TABLE 45 Baseline characteristics by presence and absence of surgical and shoulder fracture-related complications (*continued*)

Characteristic	Patients with at least one complication (<i>n</i> = 53)	Patients without complications (<i>n</i> = 197)	Total (<i>N</i> = 250)
Previous fractures, <i>n</i> (%)			
Yes	14 (26.4)	52 (26.4)	66 (26.4)
No	39 (73.6)	143 (72.6)	182 (72.8)
Missing	0 (0.0)	2 (1.0)	2 (0.8)
Previous surgery for fractures, <i>n</i> (%)			
Yes	3 (5.7)	17 (8.6)	20 (8.0)
No	10 (18.9)	34 (17.3)	44 (17.6)
Missing	1 (1.9)	1 (0.5)	2 (0.8)
No previous fractures	39 (73.6)	145 (73.6)	184 (73.6)
Injury to dominant shoulder, <i>n</i> (%)			
Yes	30 (56.6)	98 (49.7)	128 (51.2)
No	23 (43.4)	95 (48.2)	118 (47.2)
Missing	0 (0.0)	4 (2.0)	4 (1.6)

Max., maximum; min., minimum.

Treatments for surgical and shoulder fracture-related complications

Clinical staff were asked to detail any secondary surgery or increased/new therapy for these complications at the end of the orthopaedic inpatient episodes and again at the 1- and 2-year follow-ups.

After processing there were 33 valid entries reported for 33 patients, all of whom had received the treatment that they were randomised to. Each treatment corresponded to a surgical or shoulder fracture-related complication as detailed in the previous section. A breakdown of reported treatments by time and treatment group is presented in *Table 46*.

Secondary surgery to the shoulder was reported for 22 (9%) of the 250 participants, with 11 in each treatment group. Increased/new shoulder-related therapy was reported for 11 (4%) of the 250 participants, of whom seven were in the surgery group and four were in the non-surgery group. Although numbers were too small for formal testing, there appeared to be no marked group differences for secondary surgery or increased/new therapies. There was no statistically significant difference when combining secondary surgery and increased/new therapies to compare the total number of treatments between groups [$\chi^2(1) = 0.31$, $p = 0.575$].

Using all available patient information, the two expert raters assessed whether a reported surgical or shoulder fracture-related complication was addressed by further surgery or increased/new therapy and whether or not each complication was ultimately resolved (*Table 47*). Proportionally more complications were treated surgically in the non-surgery group (48% vs. 33% in the surgery group). The opposite was true for increased or new therapy (21% in the surgery group vs. 13% in the non-surgery group).

TABLE 46 Number of secondary surgeries or increased/new therapies by treatment group

Treatment for complications	Surgery (<i>n</i> = 125)	Not surgery (<i>n</i> = 125)	Total (<i>N</i> = 250)
Secondary surgery to the shoulder			
End of orthopaedic inpatient episode, <i>n</i>	0	0	0
Year 1, <i>n</i>	6	9	15
Year 2, <i>n</i>	5	2	7
Total patients, <i>n</i> (%) ^a	11 (8.8)	11 (8.8)	22 (8.8)
Increased/new shoulder-related therapy			
End of orthopaedic inpatient episode, <i>n</i>	1	0	1
Year 1, <i>n</i>	3	3	6
Year 2, <i>n</i>	3	1	4
Total patients, <i>n</i> (%) ^a	7 (5.6)	4 (3.2)	11 (4.4)
Total patients, <i>n</i> (%) ^a	18 (14.4)	15 (12.0)	33 (13.2)

^a As no multiple treatments were reported (or kept after data screening) for any patient, the number of treatments equals the number of patients.

TABLE 47 Shoulder fracture-related complications: treatments and outcomes by treatment group

Treatment for complications and outcome	Surgery (<i>n</i> = 33 complications) ^a	Not surgery (<i>n</i> = 23 complications)	Total (<i>N</i> = 56 complications)
Secondary surgery, <i>n</i> (%)			
Yes	11 (33.3)	11 (47.8)	22 (39.3)
No	22 (66.7)	12 (52.2)	34 (60.7)
No information	0 (0.0)	0 (0.0)	0 (0.0)
Increased/new therapy, <i>n</i> (%)			
Yes	7 (21.2)	3 (13.0)	10 (17.9)
No	25 (75.8)	19 (82.6)	44 (78.6)
No information	1 (3.0)	1 (4.4)	2 (3.6)
Complication outcome, <i>n</i> (%)			
Resolved	2 (6.1)	1 (4.4)	3 (5.4)
Resolving	1 (3.0)	1 (4.4)	2 (3.6)
Under review	4 (12.1)	2 (8.7)	6 (10.7)
Unresolved	2 (6.1)	2 (8.7)	4 (7.1)
Unresolved – further surgery	1 (3.0)	2 (8.7)	3 (5.4)
Not treated	6 (18.2)	4 (17.4)	10 (17.9)
No information	17 (51.5)	11 (47.8)	28 (50.0)

^a There were 36 complications in the surgery group; these included three pairs of concurrent complications that are combined here as any treatments addressed them together.

Information regarding treatment outcome was ascertainable for only half (50%) of the recorded observations, and only a small number (9%) were categorised as being resolved or resolving. There were no marked between-group differences in outcome.

The treatments and outcomes associated with the listed complications are shown for surgery in *Table 48* and for non-surgery in *Table 49*. Linking this information in this way shows, for example, that the three malunions for cross-over patients in the surgery group were 'not treated'. One patient with malunion died before further treatment could be discussed. Four of the five nonunions in non-surgical group had secondary surgery; the fifth patient died before referral for orthopaedic assessment.

Medical complications during the inpatient stay

Clinical staff were asked to select from a list of medical complications or specify any other medical complications that occurred during the patients' orthopaedic inpatient stay when applicable. End of orthopaedic inpatient episode forms were received for 111 patients in the surgery group and 23 patients in the non-surgery group.

After processing (see *Appendix 36*), there was a total of 10 medical complications for 10 patients (4% of all patients) (*Table 50*). All complications were reported for patients in the surgery group, although two of these patients were crossovers and did not receive their randomised treatment. Medical complications could be grouped into cardiac or peripheral vascular events (MI, cardiac arrest), respiratory events (two chest infections) and gastrointestinal events (norovirus, jaundice); the remainder were individual complications summarised under 'other' (urinary retention, intensive care admission and two blood transfusions).

The small number of complications limits the validity of comparing the baseline characteristics of the 10 patients with medical complications with the baseline characteristics of the 240 patients without medical complications. Nonetheless, consistent with expectations, patients experiencing an early medical complication tended to be older [mean (SD) age 71.8 (10.08) years vs. 65 (11.96) years].

Details of medical complications that occurred after the inpatient stay were collected as treatment for newly diagnosed medical complications at 1 and 2 years' follow-up (see the following section).

TABLE 48 Details of the treatments and outcomes associated with the 36 complications in the surgery group

Complication	n ^a	Recorded treatment	Outcome
Surgical site infection	2	(T: antibiotics) ^b	Resolved
		–	No information
Nerve injury	2	T: electromyography/nerve conduction tests	Resolved
		–	No information
AVN	4	–	Under review
		–	Under review
		S: revision of PHILOS implant to hemiarthroplasty	No information
		S: revision of PHILOS implant to hemiarthroplasty	Unresolved – further surgery
Implant-related complication/failure	2	–	Not treated
		T: hydrotherapy	Unresolved

continued

TABLE 48 Details of the treatments and outcomes associated with the 36 complications in the surgery group (*continued*)

Complication	n ^a	Recorded treatment	Outcome
Metalwork problems	10	S: implant (screw) removal under general anaesthetic	Unresolved
		–	Not treated
		T: treated in a U-slab as implant was 'backing out'	No information
		S: metalwork removal	No information
		–	Not treated
		S: PHILOS implant removal	No information
		S: PHILOS implant removal and rotator cuff repair	No information
		S: metalwork removal	No information
		S: removal of metalwork and rotator cuff repair ^{c1,d1}	No information
Malunion	4 ^{xxx}	S: removal of metalwork	No information
		–	Not treated ^e
		–	Not treated
		– ^{d2}	No information
Other – post-traumatic stiffness	6	T: manipulation under anaesthesia	No information
		– ^{d2}	No information
		T: physiotherapy and hydrotherapy ^{c2,d3}	No information
		S: manipulation under anaesthesia	No information
		S: manipulation under anaesthesia	No information
Other – rotator cuff tear	3	–	Under review
		(S: rotator cuff repair) ^{b,c1,d1}	No information
		(T: physiotherapy and hydrotherapy) ^{b,c2,d3}	No information
		(T: amitriptyline) ^b	Resolving
Other – complex regional pain syndrome	1		
Other – (severe) pain	1	T: pain clinic	No information
Other – unclear	1	T: investigations including nerve conduction tests for hand paraesthesia	Under review

S, surgery; T, therapy.

a Each × denotes that a crossover patient is included in the count, that is, the patient did not receive the allocated treatment.

b (T) or (S) signifies that treatment was mentioned but was not counted in the analysis as it was not recorded in the correct location on the form.

c Combined treatment (only counted once in analysis) for two separately reported complications: 1, metalwork problems and rotator cuff tear; 2, post-traumatic stiffness and rotator cuff tear.

d Concurrent complications: 1, metalwork problems and rotator cuff tear; 2, malunion and post-traumatic stiffness; 3, post-traumatic stiffness and rotator cuff tear.

e Patient died before surgery was discussed.

TABLE 49 Details of the treatments and outcomes associated with the 23 complications in the non-surgery group

Complication	n ^a	Recorded treatment	Outcome
AVN	1	S: resurfacing hemiarthroplasty	No information
Dislocation/instability	1 ^x	–	Not treated
Malunion	5	–	Not treated
		S: reverse shoulder arthroplasty	No information
		S: subacromial decompression	No information
		–	Not treated
		–	Under review
Nonunion	5	–	Not treated ^b
		S: ORIF with PHILOS plate	No information
		S: intramedullary nailing and biocomposite bone grafting	No information
		S: intramedullary nailing and calcium phosphate bone grafting	Unresolved – further surgery
		S: ORIF plus bone grafting	Unresolved
Other – post-traumatic stiffness	5	S: arthroscopic capsular release and subacromial decompression	Resolved
		T: physiotherapy	No information
		T: intensive physiotherapy	Resolving
		(T: physiotherapy) ^c	Under review
		–	No information
Other – rotator cuff tear	1	S: arthroscopic subacromial decompression plus mini-open cuff repair	Unresolved – further surgery
Other – (severe) pain	1	S: Epoca hemiarthroplasty	Unresolved
Other – impingement	1	–	No information
Other – undear	3	T: further physiotherapy	No information
		T: no information	No information
		S: PHILOS plate fixation	No information

ORIF, open reduction and internal fixation; S, surgery; T, therapy.

a Each x in this column denotes that a crossover patient is included in the count, that is, the patient did not receive the allocated treatment.

b Patient died before referral for orthopaedic assessment.

c (T) signifies that treatment was mentioned but was not counted in the analysis as it was not recorded in the correct location on the form.

TABLE 50 Number of medical complications during the inpatient stay by treatment group

Complication	Surgery (<i>n</i> = 125) ^a	Not surgery (<i>n</i> = 125)	Total (<i>N</i> = 250)
Cardiac or peripheral vascular event, <i>n</i>	2	0	2
Respiratory event, <i>n</i>	2 ^x	0	2
Gastrointestinal event, <i>n</i>	2 ^x	0	2
Other, <i>n</i>	4	0	4
Total complications, <i>n</i>	10	0	10
Total patients, <i>n</i>	10	0	10
Total patients, %	8.0	0.0	4.0
Mean complications per patient	0.08	0	0.04

a Each x denotes that a crossover patient is included in the count, that is, the patient did not receive the allocated treatment.

Other admissions and treatments: newly diagnosed medical complications and fractures

Treatments for any serious newly diagnosed medical problems were recorded at the end of the patients' inpatient episode when applicable and at the 1- and 2-year follow-ups. After processing (see *Appendix 36*), there were treatments recorded for 51 valid medical problems in 46 patients (*Table 51*).

Fifteen patients (12%) were treated for a newly diagnosed medical complication in the surgery group (including two crossover patients who did not receive surgical treatment), whereas more than twice as many patients in the non-surgery group (*n* = 31, 25%) were treated for such medical complications. However, medical complications reported for 10 patients in the surgery group during the inpatient stay (see *Table 50*) should also be taken into account.

For patients in the surgery group, the majority of treatments were recorded during the second year of follow-up, whereas in the non-surgery group an equal number of treatments were recorded in year 1 and year 2. Five patients, of whom one was in the surgery group and four were in the non-surgery group, were treated for more than one medical complication, once during the first year and once during the second year. Overall, the most common complications were 'other' (*n* = 21), cardiac and peripheral vascular events (*n* = 13, 10 of which occurred in the non-surgery group) and gastrointestinal events (*n* = 8).

Admissions to hospital for another fracture as well as any visits to the orthopaedic/fracture clinic post discharge were recorded at the 1- and 2-year follow-ups. After data processing, there were 18 valid entries for 18 patients, which are detailed in *Table 51*. Eight patients (6%) in the surgery group and 10 patients (8%) in the non-surgery group were admitted to hospital for another fracture. In addition, there was a total of 26 valid visits to orthopaedic/fracture clinics for 25 patients (see *Table 51*). Fewer patients in the surgery group (*n* = 10, 8%) than in the non-surgery group (*n* = 15, 12%) visited an orthopaedic/fracture clinic, and only one patient (non-surgery) had more than one visit for a shoulder fracture in the first year and for another fracture in the second year.

Overall, a total of 28 additional fractures for 25 patients were reported as part of the hospital admissions and clinic visits: 10 in the surgery group and 18 in the non-surgery group. More details are given in *Table 52*. The most commonly reported fractures were those of the hip (*n* = 10) and the wrist (*n* = 8). All three patients for whom multiple fractures were reported were in the non-surgical group.

TABLE 51 Numbers of other admissions and treatments by treatment group

Other admissions and treatments	Surgery (<i>n</i> = 125)				Not surgery (<i>n</i> = 125)				Total (<i>N</i> = 250)			
	Inpatient ^a	Year 1 ^b	Year 2 ^b	Total	Inpatient ^a	Year 1	Year 2	Total	Inpatient ^a	Year 1	Year 2	Total
Treatment for a newly diagnosed medical complication, <i>n</i>												
Cardiac or peripheral vascular event		2*	1	3		5	5	10		7	6	13
Respiratory event			1*	1		1	2	3		1	3	4
Gastrointestinal event			3	3		2	3	5		2	6	8
Malignancy			1	1	1	1	2	4	1	1	3	5
Other		3	5	8		8	5	13		11	10	21
Total treatments, <i>n</i>	0	5	11	16	1	17	17	35	1	22	28	51
Total patients, <i>n</i>	0	5	11	15 ^{c,d}	1	17	17	31 ^d	1	22	28	46 ^{c,d}
Total patients, %	0.0	4.0	8.8	12.0	0.8	13.6	13.6	24.8	0.4	8.8	11.2	18.4
Mean treatments per patient	0	0.04	0.0	0.13	0.01	0.14	0.14	0.28	0.0	0.09	0.11	0.20
Patient has been admitted to hospital for another fracture												
Total admissions, <i>n</i>	–	5	3	8	–	5	5	10	–	10	8	18
Total patients, <i>n</i>	–	5	3	8	–	5	5	10	–	10	8	18
Total patients, %	–	4.0	2.4	6.4	–	4.0	4.0	8.0	–	4.0	3.2	7.2
Mean admissions per patient	–	0.04	0.02	0.06	–	0.04	0.04	0.08	–	0.04	0.03	0.07

continued

TABLE 51 Numbers of other admissions and treatments by treatment group (*continued*)

Other admissions and treatments	Surgery (<i>n</i> = 125)			Not surgery (<i>n</i> = 125)			Total (<i>N</i> = 250)		
	Inpatient ^a	Year 1 ^b	Year 2 ^b	Total	Inpatient ^a	Year 1	Year 2	Total	Total
Patient has visited the orthopaedic/fracture clinic since discharge from orthopaedic treatment									
Related to primary fracture or its treatment	–	3	1	4	–	–	2	2	6
Falls and fractures at other sites	–	–	2	2	–	4	5	9	11
Other	–	1	3	4	–	1	4	5	9
Total visits, <i>n</i>	–	4	6	10	–	5	11	16	26
Total patients, <i>n</i>	–	4	6	10	–	5	11	15 ^e	25 ^e
Total patients, %	–	3.2	4.8	8.0	–	4.0	8.8	12.0	10.0
Mean visits per patient	–	0.03	0.05	0.08	–	0.04	0.09	0.13	0.10

^a At the end of the orthopaedic inpatient episode.

^b Each x in this column denotes that a crossover patient is included in the count, that is, the patient did not receive the allocated treatment.

^c One patient in the surgery group was treated for a newly diagnosed medical complication in year 1 and year 2 (other + cardiac or peripheral vascular).

^d Four patients in the non-surgery group were treated for a newly diagnosed medical complication in year 1 and year 2 (respiratory + malignancy; other + other; cardiac or peripheral vascular + gastrointestinal; cardiac or peripheral vascular + cardiac or peripheral vascular).

^e One patient in the non-surgery group visited the orthopaedic/fracture clinic in year 1 and in year 2 (shoulder fracture + other fracture).

TABLE 52 Further fractures by treatment group

Fracture site	Surgery (<i>n</i> = 125)	Not surgery (<i>n</i> = 125)	Total (<i>N</i> = 250)
Shoulder, <i>n</i>	1	2	3
Wrist, <i>n</i>	3	5	8
Hip, <i>n</i>	3	7	10
Other, <i>n</i>	3	4	7
Total fractures, <i>n</i>	10	18	28
Total patients, <i>n</i> (%)	10 (8.0)	15 (12.0) ^a	25 (10.0) ^a

a Three patients in the non-surgery group had multiple additional fractures (two in each case): hip fracture + other fracture in the first year; hip fracture + wrist fracture in the second year; shoulder fracture in the first year + other fracture in the second year.

Mortality

Of the 14 patients who died during the trial period, nine (7%) were in the surgery group and five (4%) were in the non-surgery group. Causes of death by treatment group are presented in *Table 53*. One death in the surgery group was judged as being trial related (venous thromboembolism). A chi-squared test revealed that there was no statistically significant difference in mortality rate between treatment groups [$\chi^2(1) = 1.21$, $p = 0.271$].

Adverse events

A total of 78 SAEs were reported for 56 patients and 10 non-SAEs were reported for nine patients.

Serious adverse events

Twenty-eight patients in each treatment group experienced at least one SAE, with one patient in the non-surgery group experiencing seven SAEs (*Table 54*). The numbers of SAEs by category and treatment group are given in *Table 55*. There were no obvious differences between treatment groups except that there were a greater number of serious non-fatal medical complications in the non-surgery group than in the surgery group (27 vs. 11). In part, these reflect multiple recurrences of related conditions in patients

TABLE 53 Mortality by treatment group

Cause of death	Surgery (<i>n</i> = 125)	Not surgery (<i>n</i> = 125)	Total (<i>N</i> = 250)
Lung cancer	2	1	3
Respiratory failure	2	1	3
Venous thromboembolism	1	0	1
Pneumonia	0	1	1
MI	1	1	2
Renal failure	1	0	1
Intracerebral haemorrhage	0	1	1
Accidental death	1	0	1
Unknown cause	1	0	1
Total deaths, <i>n</i> (%)	9 (7.2)	5 (4.0)	14 (5.6)

TABLE 54 Numbers of patients experiencing one or more SAE by treatment group

Number of SAEs	Surgery (<i>n</i> = 125), <i>n</i> (%)	Not surgery (<i>n</i> = 125), <i>n</i> (%)	Total (<i>N</i> = 250), <i>N</i> (%)
1	23 (18.4)	20 (16.0)	43 (17.2)
2	3 (2.4)	6 (4.8)	9 (3.6)
3	2 (1.6)	0 (0.0)	2 (0.8)
4	0 (0.0)	1 (0.8)	1 (0.4)
7	0 (0.0)	1 (0.8)	1 (0.4)
Total patients	28 (22.4)	28 (22.4)	56 (22.4)

TABLE 55 Numbers of SAEs by category and treatment group^a

SAE type	Surgery (<i>n</i> = 125) (35 events), <i>n</i> (%)	Not surgery (<i>n</i> = 125) (43 events), <i>n</i> (%)	Total (<i>N</i> = 250) (78 events), <i>N</i> (%)
Surgery to the shoulder (any mention of surgery to the shoulder)	8 (22.9)	7 (16.3)	15 (19.2)
Non-surgical treatment/ investigation of the shoulder	1 (2.9)	1 (2.3)	2 (2.6)
Medical complication/ condition			
Non-serious	1 (2.9)	0 (0.0)	1 (1.3)
Serious: non-fatal	11 (31.4)	27 (62.8)	38 (48.7)
Serious: fatal	9 (25.7)	5 (11.6)	14 (17.9)
Any fracture			
Shoulder	0 (0.0)	1 (2.3)	1 (1.3)
Wrist	0 (0.0)	0 (0.0)	0 (0.0)
Hip	3 (8.6)	2 (4.7)	5 (6.4)
Other	2 (5.7)	0 (0.0)	2 (2.6)

^a Percentages relate to the total number of SAEs in each treatment group and overall.

experiencing several events but there is also a parallel with the greater number of patients treated for serious newly diagnosed medical problems in the non-surgery group, recorded on the hospital forms (see *Table 51*).

Non-serious adverse events

Nine patients experienced at least one non-SAE, of whom four were in the surgery group and five were in the non-surgery group. The numbers of patients experiencing one or more non-SAE are presented by treatment group in *Table 56*. The numbers of non-SAEs by category and treatment group are given in *Table 57*. With so few non-SAEs reported there were no obvious differences between treatment groups.

TABLE 56 Numbers of patients experiencing one or more non-SAE by treatment group

Number of non-SAEs	Surgery (<i>n</i> = 125), <i>n</i> (%)	Not surgery (<i>n</i> = 125), <i>n</i> (%)	Total (<i>N</i> = 250), <i>N</i> (%)
1	4 (3.2)	4 (3.2)	8 (3.2)
2	0 (0.0)	1 (0.8)	1 (0.4)
Total	4 (3.2)	5 (4.0)	9 (3.6)

TABLE 57 Numbers of non-SAEs by category and treatment group^a

Non-SAE type	Surgery (<i>n</i> = 125) (four events), <i>n</i> (%)	Not surgery (<i>n</i> = 125) (six events), <i>n</i> (%)	Total (<i>N</i> = 250) (10 events), <i>N</i> (%)
Non-surgical treatment/ investigation of the shoulder	1 (25.0)	0 (0.0)	1 (10.0)
Medical complication/condition			
Non-serious	1 (25.0)	5 (83.3)	6 (60.0)
Serious: non-fatal	1 (25.0)	0 (0.0)	1 (10.0)
Any fracture: wrist	1 (25.0)	1 (16.7)	2 (20.0)

^a Percentages relate to the total number of SAEs in each treatment group and overall.

Expected adverse events and relatedness of adverse events to taking part in the trial

Expected adverse events and trial-related adverse events (serious and non-serious) as reported by local investigators were more likely to occur in the surgery group than in the non-surgery group. This is likely to reflect the fact that it is simpler to link adverse events to surgery in the surgery group than it is to link them to the treatment received in the non-surgery group (*Table 58*).

TABLE 58 Expected adverse events and relatedness of adverse events to trial participation by treatment group^a

Type of adverse event	Surgery (<i>n</i> = 125) (39 events), <i>n</i> (%)	Not surgery (<i>n</i> = 125) (49 events), <i>n</i> (%)	Total (<i>N</i> = 250) (88 events), <i>N</i> (%)
Expected			
Yes	11 (28.2)	5 (10.2)	16 (18.2)
No	28 (71.8)	44 (89.8)	72 (81.8)
Related to trial			
Yes	9 (23.1)	2 (4.1)	11 (12.5)
No	30 (76.9)	47 (95.9)	77 (87.5)
Combinations			
Unrelated to trial and expected	6 (15.4)	4 (8.2)	10 (11.4)
Unrelated to trial and unexpected	24 (61.5)	43 (87.8)	67 (76.1)
Related to trial and expected	5 (12.8)	1 (2.0)	6 (6.8)
Related to trial and unexpected	4 (10.3)	1 (2.0)	5 (5.7)

^a Percentages relate to the total number of adverse events in each group and overall.

'Untoward events' documented in the physiotherapy treatment logs

The screening by two independent raters (LG and HH) of comments and reasons for referral recorded in the physiotherapy treatment logs resulted in the identification of 37 'untoward events' recorded for 32 patients. These were shoulder-related or medical events that impeded physiotherapy or were reasons for non-attendance or for referral. Listed surgical or shoulder-related complications that were not identified by the inpatient and two hospital follow-up forms are summarised in *Tables 59* and *60* for the surgical group and non-surgical group respectively. As in the coding for the hospital forms, frozen shoulder is termed post-traumatic stiffness. In screening of the 'untoward events', precursors (e.g. severe pain) to a complication recorded on the hospital form (e.g. AVN) and complications subsequent to treatment for a complication were not included.

Given that the physiotherapy logs did not specifically collect data on complications and considering the diffuseness and variability in the commentary, these results should be considered as illustrative of what was recorded. It is clear that a few early complications were not recorded on the hospital forms but, apart from axillary nerve lesion, none of these complications was a named complication on the hospital forms.

TABLE 59 Additional complications reported in the physiotherapy logs for surgical group patients

Patient ID ^a	Complication	Physiotherapy logs: other information	Hospital forms: information
S01	Delayed union	Protracted physiotherapy	Healed eventually
S02	Chronic pain and allodynia of the shoulder	Problematic recovery because of pain	No mention; developed hand paraesthesia
S03	Complex regional pain syndrome (elements of)	Advised to see GP to obtain amitriptyline (for neuropathic pain); resolving	No mention
S04	Axillary nerve lesion	Referral – orthopaedics; recovered spontaneously	No mention
S05	Shoulder pain and swelling – metalwork removal considered?	Referral – orthopaedics; underwent CT scan	No mention
S06	Rotator cuff tear	No mention	No mention; post-traumatic stiffness

^a Patient IDs are specific to this table.

TABLE 60 Additional complications reported in the physiotherapy logs for non-surgical group patients

Patient ID ^a	Complication	Physiotherapy logs: other information	Hospital forms: information
NS01	Humeral head subluxation	Referral – orthopaedics	No mention; post-traumatic stiffness
NS02	Post-operative stiffness	Advised to see GP to obtain pain relief	No mention; developed AVN
NS03	Patient felt a 'crack' and pain whilst exercising	Given local anaesthesia injection	No mention; eventual nonunion
	Post-traumatic stiffness	No mention	No mention
NS04	Delayed union (shoulder pain)	Referral – orthopaedics	Healed eventually
NS05	Rotator cuff tear developing	No mention	No mention

^a Patient IDs are specific to this table.

Discharge locations after inpatient stay

Only five of the 109 participants who were inpatients in the surgery group and one of the 22 participants who were inpatients in the non-surgery group were discharged to locations other than back home. In the surgery group, one patient was discharged to a nursing home, three were discharged to a rehabilitation unit and one was discharged to a renal ward for dialysis. The patient in the non-surgery group was discharged to another hospital.

Patient treatment preferences at the end of the 2-year follow-up

Consenting patients were asked at baseline in advance of randomisation whether or not they had any treatment preference before agreeing to randomisation. At the end of the 2-year follow-up period, patients were again asked to indicate their treatment preference in the event of a similar shoulder injury given their experience over the past 2 years.

Patient preferences at 2 years' follow-up are tabulated against baseline preferences in *Table 61* and against the allocated treatment in *Table 62*. More patients expressed a preference for a treatment (either surgery or not surgery) at 24 months (61.9%) than at baseline (53.4%). The slight preference for surgical treatment over non-surgical treatment seen at baseline remained.

Fewer than half of the patients retained their original treatment preference (41.7% for surgery and 46.7% for non-surgery) (see *Table 61*). Around half of the patients expressed a preference for the treatment that they were allocated to (51.2% in the surgery group and 44.0% in the non-surgery group) (see *Table 62*). In each case, patients' second most frequent choice was 'no preference' rather than the opposite treatment.

TABLE 61 Treatment preferences of trial participants at baseline and 24 months

Patient preference at 24 months	Patient preference at baseline, <i>n</i> (%)		
	Surgery (<i>n</i> = 72)	Not surgery (<i>n</i> = 60)	No preference (<i>n</i> = 115)
Surgery	30 (41.7)	15 (25.0)	39 (33.9)
Not surgery	16 (22.2)	28 (46.7)	25 (21.7)
No preference	16 (22.2)	9 (15.0)	29 (25.2)
Missing	10 (13.9)	8 (13.3)	22 (19.1)

TABLE 62 Treatment preferences of trial participants at 24 months by treatment group

Patient preference at 24 months	Randomised group, <i>n</i> (%)		
	Surgery (<i>n</i> = 125)	Not surgery (<i>n</i> = 125)	Total (<i>N</i> = 250)
Surgery	64 (51.2)	22 (17.6)	86 (34.4)
Not surgery	15 (12.0)	55 (44.0)	70 (28.0)
No preference	24 (19.2)	30 (24.0)	54 (21.6)
Missing	22 (17.6)	18 (14.4)	40 (16.0)

Lack of surgeon equipoise

As reported in *Chapter 4*, the independent assessment of the 'other' reasons given for non-eligibility on the study eligibility form (option j; see *Appendix 5*) resulted in the identification of 117 patients who were judged to have been excluded because of a lack of equipoise. The two raters also recorded the treatment preference stated for individual patients (surgery, conservative management, unclear) and deduced this when not recorded in the appropriate box on the study eligibility form if other information was available.

Listed below is the categorisation of 'protocol violations' for the 159 study eligibility forms on which only option j had been ticked:

- lack of clinical equipoise: 117 (74%) [surgeons' treatment decision/preference: surgery 45 (38%); not surgery 63 (54%); unclear 9 (8%)]
- existing exclusion criteria: 18 (11%)
- local difficulties (e.g. research nurse not available): 11 (7%)
- new exclusion criteria: 8 (5%)
- other: 5 (3%).

Listed under 'new exclusion criteria', which equated to when the two raters considered that the reason for exclusion was justified, were the following:

- severely compromised social circumstances ($n = 1$)
- repeat injury to previous fracture ($n = 1$)
- extreme age and frailty ($n = 1$)
- patient with compromised or exceptional use (e.g. wheelchair user) upper limb function ($n = 4$)
- head-split fracture ($n = 1$).

Between one and 23 patients were excluded because of a lack of equipoise in 22 centres; no patients were excluded for this reason in the remaining 11 centres that returned study eligibility forms. *Figure 21* shows the numbers of patients excluded because of a lack of equipoise together with the numbers of ineligible patients for each of the 33 centres. Lack of equipoise is most notable in three centres, two with high numbers of inappropriately excluded patients [23 (47% of ineligible patients) and 14 (67% of ineligible patients)] and one in which lack of equipoise was the reason for exclusion of all seven ineligible patients.

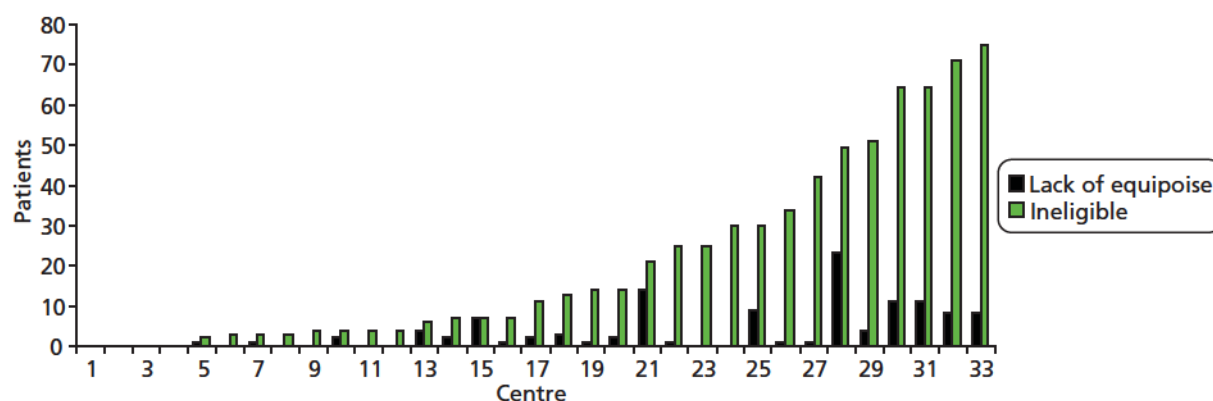


FIGURE 21 Lack of surgeon equipoise and ineligible patients by centre.

Lack of equipoise was recorded for 68 surgeons. In some centres only a small proportion of surgeons (e.g. one of 16 in one centre) listed on study eligibility forms excluded patients because of a lack of equipoise, whereas in other centres this was more prevalent (e.g. 12 of 17 surgeons in one centre). Several surgeons ($n = 12$) who showed a lack of equipoise screened and excluded only one patient. Of those who excluded more than one patient because of a lack of equipoise, a further 12 excluded half or more of the patients who they screened.

Using the t -test or chi-squared test as applicable, selected baseline characteristics were compared between all eligible patients and patients who were excluded because of a lack of equipoise to identify any differences between the two populations (*Table 63*). Patients who were excluded because of a lack of equipoise tended to be slightly older (70 years vs. 67 years, $p = 0.062$) and were more likely to have a fracture not involving either tuberosity (41% vs. 25%, $p < 0.001$) than the sample of all eligible patients. The proportion of men and women as well as time since injury did not appear to differ meaningfully between the two groups.

Age and tuberosity involvement were tabulated against surgeons' treatment preference (reported data only) and assessed by chi-squared test for patients excluded because of a lack of equipoise to estimate the direction and strength of any observed preference related to age and fracture type (*Table 64*). Patients aged ≥ 65 years were more likely to be advised to have non-surgical treatment ($p = 0.059$) whereas patients with fractures involving tuberosities were just as likely to be advised surgery or non-surgery as patients with fractures involving no tuberosities ($p = 0.794$).

TABLE 63 Baseline characteristics by lack of equipoise exclusion

Characteristic	Total eligible patients (<i>n</i> = 563)	Patients excluded because of a lack of equipoise (<i>n</i> = 117)	Significance of group difference	
			Test	<i>p</i> -value
Age (years)				
<i>n</i>	548	113	<i>t</i> (659) = −1.9	0.062
Mean (SD)	67.4 (12.60)	69.9 (14.63)		
Median (min., max.)	69.1 (18.4, 98.4)	72.0 (16.1, 106.2)		
Gender, <i>n</i> (%)				
Male	133 (23.6)	25 (21.4)	$\chi^2(1)$ = 0.3	0.586
Female	428 (76.0)	92 (78.6)		
Missing	2 (0.4)	0		
Time since injury (days)				
<i>n</i>	563	117	<i>t</i> (678) = −0.3	0.741
Mean (SD)	5.4 (4.80)	5.6 (5.41)		
Median (min., max.)	4 (0, 21)	3 (0, 19)		
Tuberosity involvement, <i>n</i> (%)				
Yes	422 (75.0)	69 (59.0)	$\chi^2(1)$ = 12.3	0.000
No	141 (25.0)	48 (41.0)		
Max., maximum; min., minimum.				

TABLE 64 Baseline characteristics by surgeon advice for patients excluded because of a lack of equipoise

Characteristic	Surgeon preference: surgery (<i>n</i> = 41)	Surgeon preference: not surgery (<i>n</i> = 58)	Significance of group difference	
			Test	<i>p</i> -value
Age (years), <i>n</i> (%)				
< 65	16 (39.0)	12 (20.7)	$\chi^2(1) = 3.6$	0.059
≥ 65	25 (61.0)	44 (75.9)		
Missing	0	2 (3.4)		
Tuberosity involvement, <i>n</i> (%)				
Yes	23 (56.1)	31 (53.4)	$\chi^2(1) = 0.1$	0.794
No	18 (43.9)	27 (46.6)		

Surgeon preferences and agreed treatments for non-consenting patients

Surgeon preferences

Preferences in terms of advised treatment as judged by the recruiting clinicians were recorded for all patients who were ineligible or who did not consent to participate in the trial. The resulting frequencies in *Table 65* show that, of the ineligible patients, 34% were advised to have surgical treatment and 56% were advised to have non-surgical treatment. In contrast, only 21% of non-consenters were advised to have surgical treatment, with 34% advised to have non-surgical treatment and clinicians uncertain what treatment to advise in 38%.

Agreed treatment for non-consenting patients

Agreed treatments for patients who were eligible but who did not consent to partake in the trial are presented in *Table 66* by patient and clinician preference. A total of 55 patients expressed a preference for surgery and this treatment was agreed for 50 (91%); in two cases (4%) this was contrary to clinician advice for the patient to undergo non-surgical treatment. A total of 226 patients expressed a preference for non-surgery and this treatment was agreed for 221 (98%); in 26 cases (12%) this was contrary to clinician advice for the patient to undergo surgical intervention. Of the 23 patients without a treatment preference, surgery was agreed for seven (30%) and non-surgical treatment was agreed for 16 (70%).

TABLE 65 Treatment advice by clinicians for patients who were not entered into the trial

Clinician treatment advice	Ineligible (n = 687), n (%)	Non-consent (n = 313), n (%)
Surgery	232 (33.8)	66 (21.1)
Not surgery	384 (55.9)	105 (33.5)
Uncertain ^a	–	118 (37.7)
Missing data	71 (10.3)	24 (7.7)

^a There was no tick box for 'uncertain' on the study eligibility form. The box for 'uncertain' on the consent status form was added in response to an early detected missing data issue.

TABLE 66 Preferences and agreed treatments for eligible patients who did not consent

Clinician advice	Agreed treatment	Patient preference, n (%)			
		Surgery (n = 55)	Not surgery (n = 226)	No preference (n = 23)	Missing (n = 9)
Surgery	Surgery	27 (49.1)	3 (1.3)	6 (26.1)	0 (0.0)
	No surgery	0 (0.0)	26 (11.5)	3 (13.0)	0 (0.0)
	Missing	1 (1.8)	0 (0.0)	0 (0.0)	0 (0.0)
No surgery	Surgery	2 (3.6)	0 (0.0)	0 (0.0)	0 (0.0)
	No surgery	4 (7.3)	94 (41.6)	4 (17.4)	1 (11.1)
	Missing	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Uncertain	Surgery	18 (32.7)	0 (0.0)	1 (4.3)	0 (0.0)
	No surgery	0 (0.0)	87 (38.5)	9 (39.1)	0 (0.0)
	Missing	0 (0.0)	2 (0.9)	0 (0.0)	1 (11.1)
Missing	Surgery	3 (5.5)	0 (0.0)	0 (0.0)	0 (0.0)
	No surgery	0 (0.0)	14 (6.2)	0 (0.0)	0 (0.0)
	Missing	0 (0.0)	0 (0.0)	0 (0.0)	7 (77.8)

Availability of radiographic views in the study population

Clinicians completing a study eligibility form (see *Appendix 5*) recorded whether or not the series of views used to inform the assessment of the fracture included one or more of the following views: the anteroposterior view, the axillary (or modified axillary) view and the scapular Y-lateral view. Our protocol stipulated that a minimum of two radiographic views/projections were required for the assessment of study eligibility. *Table 67* shows the breakdown of radiographic views recorded for all screened patients. This demonstrates that the majority (88%) of the study eligibility forms indicated that the minimum requirement was achieved, with the standard trauma series of three perpendicular views⁵⁹ being available for 18% of screened patients. The anteroposterior view was recorded in 98% of patients, the axillary view in 57% and the scapular Y-lateral view in 50%. Similar findings in terms of views recorded and meeting the minimum requirements of two perpendicular views applied for ineligible (88%), eligible (89%) and randomised patients (89%).

TABLE 67 Radiographic views recorded on the study eligibility forms

Radiographic views	Screened (n = 1250), n (%)	Ineligible (n = 687), n (%)	Eligible (n = 563), n (%)	Randomised (n = 250), n (%)
Anteroposterior view only	117 (9.4)	64 (9.3)	53 (9.4)	22 (8.8)
Axillary view only	2 (0.2)	1 (0.1)	1 (0.2)	0 (0.0)
Scapular Y-lateral view only	1 (0.1)	1 (0.1)	0 (0.0)	0 (0.0)
Anteroposterior + axillary views ^a	480 (38.4)	284 (41.3)	196 (34.8)	85 (34.0)
Anteroposterior + scapular Y-lateral views ^a	398 (31.8)	220 (32.0)	178 (31.6)	61 (24.4)
Axillary + scapular Y-lateral views ^a	2 (0.2)	2 (0.3)	0 (0.0)	0 (0.0)
Anteroposterior + axillary + scapular Y-lateral views ^a	224 (17.9)	98 (14.3)	126 (22.4)	76 (30.4)
Missing	26 (2.1)	17 (2.5)	9 (1.6)	6 (2.4)

a Two or three views complied with the requirements for assessing study eligibility.

Use of other imaging investigations

Data on the use of other imaging modalities to determine study eligibility were collected only for randomised patients (see *Appendix 20*). In total, six patients (three in each treatment allocation group), each from a different hospital, received a CT scan. This is consistent with the results of the post-recruitment survey of radiography departments, which found that CT scanning was 'routinely/frequently used for these fractures' in only one of the 26 responding hospitals (see *Appendix 31*).

Independent assessment and classification of baseline radiographs according to the Neer classification

Cull of radiographs in preparation for assessment

A minimum of two radiographs were available for each trial participant, with a maximum of seven being received for two participants. In all, 664 radiographs were available for the 250 study participants. Sixty-two views were excluded based on prespecified criteria, which aimed to retain the best-quality view(s) in each perpendicular plane. Thus, the number of named views in each plane would not have been affected by this cull. The results of the cull are illustrated in *Figure 22*. For the majority of patients ($n = 165$), two radiographic views were available post cull, leaving 68 with three views and 17 with four views.

Assessment of available radiographic views

The two independent surgeons (raters), using the Neer classification proforma (see *Appendix 33*), were asked to indicate which views were available to them (*Table 68*). Compared with the baseline assessment, raters judged that there were a greater number of anteroposterior plus scapular Y-lateral views and fewer anteroposterior only views and anteroposterior plus axillary plus scapular Y-lateral views. There was little difference in judgement between the two raters (91% agreement, kappa = 0.87, $p < 0.001$), whereas agreement compared with baseline was weaker (rater 1: 56% agreement, kappa = 0.39, $p < 0.001$; rater 2: 59% agreement, kappa = 0.43, $p < 0.001$). It is noteworthy that the 10 views for which either rater indicated a single plane only (excluding any additional 'other' views) had all been rated to show a minimum of two planes on the eligibility form.

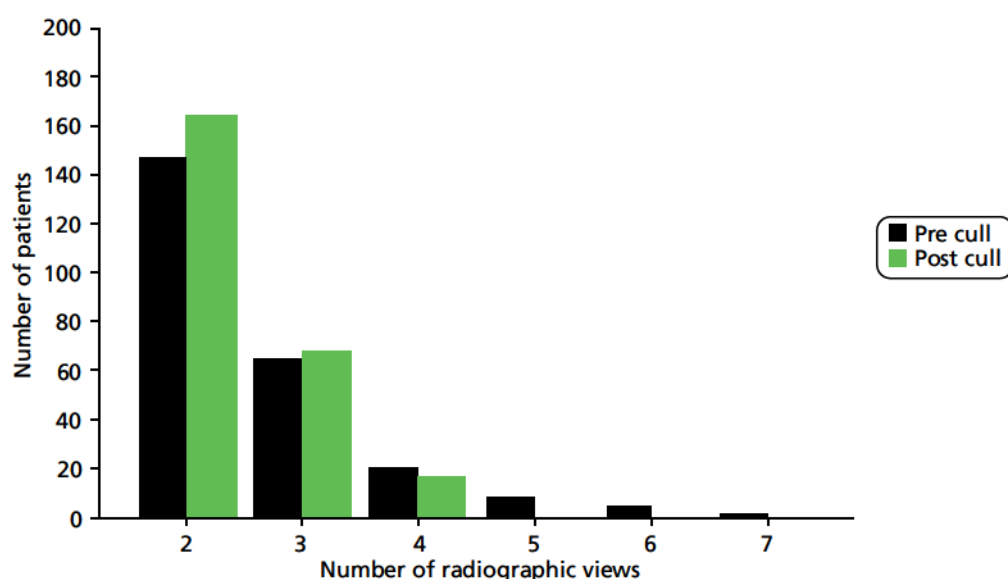


FIGURE 22 Pre- and post-cull radiographic views per patient.

TABLE 68 Radiographic views available for study participants (eligibility forms and Neer classification proforma)

Radiographic view (eligibility forms and Neer classification proforma)	Eligibility form (<i>n</i> = 250), <i>n</i> (%)	Rater 1 (<i>n</i> = 250), <i>n</i> ^b (%)	Rater 2 (<i>n</i> = 250), <i>n</i> ^b (%)
Anteroposterior only ^a	22 (8.8)	10 ¹ (4.0)	7 ¹ (2.8)
Axillary only	0 (0.0)	1 ¹ (0.4)	0 (0.0)
Scapular Y-lateral only	0 (0.0)	1 ¹ (0.4)	0 (0.0)
Anteroposterior + axillary	85 (34.0)	87 ⁴ (34.8)	82 (32.8)
Anteroposterior + scapular Y-lateral	61 (24.4)	91 ³ (36.4)	95 (38.0)
Axillary + scapular Y-lateral	0 (0.0)	1 ¹ (0.4)	1 (0.4)
Anteroposterior + axillary + scapular Y-lateral	76 (30.4)	58 ¹ (23.2)	65 ² (26.0)
Missing	6 (2.4)	1 (0.4)	0 (0.0)

a Combines scapular and coronal plane views (Neer classification only).

b Superscript numbers (1–4) indicate how many patients had the named radiographic view(s) plus one 'other' view (Neer classification only).

Assessment of radiograph quality

As part of the completion of the Neer classification proforma, the two independent raters were asked to evaluate the quality of the radiographs available for each patient in terms of their adequacy for classification purposes. The questions asked are detailed in Table 69. A number of responses to questions 2 and 4 were missing for both raters but could be inferred from the individual assessments of each radiograph that were provided. For question 2, 14 responses were inferred for rater 1 and 35 for rater 2. For question 4, five responses were inferred for rater 1 and six for rater 2.

TABLE 69 Agreement on radiograph quality criteria between raters

Question (from the Neer classification proforma)	Response	Rater 1 (<i>n</i> = 250), <i>n</i> (%)	Rater 2 (<i>n</i> = 250), <i>n</i> (%)	Agreement between raters ^a
Q2. Are there at least two projections in planes perpendicular to each other?	Yes	239 (95.6)	241 (96.4)	97.6% agreement (kappa = 0.69, <i>p</i> < 0.001)
	No	11 (4.4)	9 (3.6)	
Q4. Are the proximal humerus and glenohumeral joint seen on each projection (of at least two different views)?	Yes	245 (98.0)	241 (96.4)	95.2% agreement (kappa = 0.12, <i>p</i> = 0.023)
	No	5 (2.0)	9 (3.6)	
Q6. Considering all the available views together, can you visualise the location of all five structures (the humeral shaft, greater tuberosity, lesser tuberosity, head of the humerus and glenohumeral joint) sufficiently to determine the position and displacement of the fractured segments?	Yes	248 (99.2)	234 (93.6)	93.6% agreement (kappa = 0.10, <i>p</i> = 0.006)
	No	2 (0.8)	16 (6.4)	
Q7. Please give your overall assessment of quality for this set of images (i.e. in terms of exposure and patient positioning) based on the three grades good, fair and poor	Good	219 (87.6)	168 (67.2)	71.4% agreement (kappa = 0.23, <i>p</i> < 0.001)
	Fair	29 (11.6)	66 (26.4)	
	Poor	2 (0.8)	14 (5.6)	
	Missing	0 (0.0)	2 (0.8)	

a Although, as planned, kappa statistic results are presented, these serve mainly to illustrate that kappa statistics are misleading when there is a lack of variability in the responses.

Overall, *Table 69* shows that the sets of radiographs provided perpendicular views (for 96% of patients) that were sufficient to see the relevant anatomical structures (96–98%) and visualise the location of these (94–99%). Agreement on these questions was high between raters (94–98%). Some of the associated kappa values are low, which can be explained by the lack of variability in the responses, which results in a high level of expected chance agreement against which the actual agreement is compared when using the kappa statistic. Hence, the percentage agreement is the more representative statistic in this table.

Rater 2 judged the overall quality of the images (question 7) to be poorer than rater 1. Rater 1 estimated that 88% of the radiograph sets were of good quality whereas rater 2 estimated that only 67% of the radiograph sets were of good quality. Accordingly, the level of agreement between raters was lower for this question (71% agreement, kappa = 0.23, $p < 0.001$).

Fracture types and characteristics: agreement between the two raters

Using assessments of individual anatomical features for each radiograph and arriving at a verdict for each feature's involvement/displacement, the two raters independently assigned a Neer category for each patient (see *Appendix 2* for the 16 possible categories). The classifications assigned in the first stage of the classification process by each rater are listed in *Table 70*.

The majority of fractures were judged to be two-part surgical neck fractures (category 3, 51–52% of patients) and three-part surgical neck plus greater tuberosity fractures (category 8, 32%). Rater 2 was more likely to assign fractures to category 1 (single-part fractures) and category 4 (two-part greater tuberosity fractures) whereas rater 1 was more likely to assign fractures to category 12 (four-part fractures involving the surgical neck and both tuberosities). Agreement between the two raters was therefore only moderate (68% agreement, kappa = 0.48, $p < 0.001$); overall, the two raters independently assigned the same category in 169 cases.

The presence of other individual characteristics listed on the Neer classification proforma as judged by the two raters is given in *Table 71*. In total, 9–11% of the fractures were judged to be no contact surgical neck fractures and 24–30% were judged to be impacted surgical neck fractures. There was one dislocation and 4% were articular surface fractures. The head segment was in varus in 20–25% of fractures and in valgus in 28–31% of fractures. Differences between raters applied to only a small number of these features, with a maximum magnitude of difference of 7%, which was for impaction.

TABLE 70 Neer classification: initial assignments by the two raters acting independently

Neer classification	Rater 1 ($n = 250$), n (%)	Rater 2 ($n = 250$), n (%)
1 Neer one part: undisplaced ^a surgical neck	8 (3.2)	21 (8.4)
3 Neer two part: surgical neck	127 (50.8)	129 (51.6)
4 Neer two part: greater tuberosity ^b	2 (0.8)	14 (5.6)
5 Neer two part: lesser tuberosity ^b	1 (0.4)	0 (0.0)
8 Neer three part: surgical neck + greater tuberosity	81 (32.4)	81 (32.4)
9 Neer three part: surgical neck + lesser tuberosity	0 (0.0)	0 (0.0)
10 Neer three part: anterior dislocation + greater tuberosity	0 (0.0)	0 (0.0)
12 Neer four part: surgical neck + greater tuberosity + lesser tuberosity	30 (12.0)	4 (1.6)
13 Fracture–dislocation – anterior (four part)	1 (0.4)	0 (0.0)
15 Fracture–dislocation – anterior (articular surface)	0 (0.0)	1 (0.4)

a Fractures were rated as 'undisplaced' for the purposes of the Neer classification but were in reality displaced fractures in which the degree of displacement was insufficient for categorisation as a 'part'.

b Surgical neck involved but not sufficiently displaced to be a 'part'.

TABLE 71 Separate fracture characteristics on the Neer classifications proforma: individual judgements of raters

Fracture characteristic	Rater 1 (n = 250), n (%)	Rater 2 (n = 250), n (%)
Surgical neck – no contact fracture	23 (9.2)	27 (10.8)
Surgical neck – impacted	59 (23.6)	76 (30.4)
Anterior fracture–dislocation	1 (0.4)	0 (0.0)
Posterior fracture–dislocation	0 (0.0)	0 (0.0)
Articular surface fracture	11 (4.4)	10 (4.0)
Head segment in varus	62 (24.8)	49 (19.6)
Head segment in valgus	69 (27.6)	77 (30.8)

Agreed Neer classification of baseline fractures

Consensus was reached between the two raters for the 81 fractures for which there had been discordance in the initial assignment of Neer classification. The agreed classifications are provided in *Table 72* by trial arm. The majority of patients were still assigned to category 3 (48%) and category 8 (36%); however, entries in categories 13 and 15 were removed from the individual assessments and added to categories 9 and 10 instead. Classifications were very well balanced between the trial arms.

As highlighted in *Table 72*, some categories were ‘unexpected’ fractures (categories 1, 4, 5 and 10), particularly those that were not associated with the displacement of the surgical neck (Neer criteria). *Tables 73* and *74* provide further details for fractures assigned to these categories.

For all categories representing an ‘undisplaced’ surgical neck (category 1: *n* = 18 undisplaced surgical neck only; category 4: *n* = 8 undisplaced surgical neck plus greater tuberosity; category 5: *n* = 1 undisplaced surgical neck plus lesser tuberosity), the surgical neck was assessed to be involved in all cases by both raters, with the majority indicating displacement (Neer or unclear Neer: 74% rater 1, 95% rater 2). For most of the category 1 and category 4 fractures (85–88%), the greater tuberosity was also involved, whereas both tuberosities were estimated to be involved for the one category 5 fracture. *Table 74* shows that this agrees with data from the study eligibility forms, confirming greater tuberosity involvement (either on its own or together with lesser tuberosity involvement) for the majority of Neer category 1 (67% + 14%) and category 4 (94% + 6%) fractures and the involvement of both tuberosities for category 5 fractures (100%).

TABLE 72 Agreed Neer classification by treatment group

Final Neer classification categories (see <i>Appendix 2</i>) ^a	Surgery (n = 125), n (%)	Not surgery (n = 125), n (%)
1 ^b Neer one part: undisplaced ^c surgical neck	9 (7.2)	9 (7.2)
3 Neer two part: surgical neck	60 (48.0)	59 (47.2)
4 ^b Neer two part: greater tuberosity	5 (4.0)	3 (2.4)
5 ^b Neer two part: lesser tuberosity	0 (0.0)	1 (0.8)
8 Neer three part: surgical neck + greater tuberosity	45 (36.0)	45 (36.0)
9 Neer three part: surgical neck + lesser tuberosity	1 (0.8)	0 (0.0)
10 ^b Neer three part: anterior dislocation + greater tuberosity	0 (0.0)	2 (1.6)
12 Neer four part: surgical neck + greater tuberosity + lesser tuberosity	5 (4.0)	6 (4.8)

^a The categories in bold (categories 3, 8, 9 and 12) are those that fully meet the ProFHER trial inclusion criteria.

^b These are categories outside the expected categories for the trial.

^c Fractures were rated as ‘undisplaced’ for the purposes of the Neer classification but were in reality displaced fractures in which the degree of displacement was insufficient for categorisation as a ‘part’.

TABLE 73 'Unexpected' Neer categories: involvement/displacement of anatomical structures (Neer classification proforma)

Neer classification	Rater 1		Rater 2			
	Anatomical structure	Involvement	Undisplaced	Displaced (unclear if Neer displacement criteria met)	Displaced (Neer)	Displaced (Neer)
1	Surgical neck	18	4	9	5	6
	Greater tuberosity	15	5	3	7	1
	Lesser tuberosity	2	1	0	1	0
4	Surgical neck	8	2	5	1	6
	Greater tuberosity	8	0	1	7	6
	Lesser tuberosity	0	0	0	0	1
5	Surgical neck	1	1	0	0	0
	Greater tuberosity	1	0	1	0	0
	Lesser tuberosity	1	1	0	0	1
10	Surgical neck	2	0	0	2	2
	Greater tuberosity	2	0	0	2	2
	Lesser tuberosity	1	0	0	1	0

TABLE 74 'Unexpected' Neer categories: involvement of the tuberosities (Neer classification proforma plus eligibility forms)

Neer classification (either rater) ^a		Study eligibility form: tuberosity involvement			
Category	Tuberosity involvement	None	Greater	Lesser	Both
1	None	4	2	–	1
	Greater	–	19	–	5
	Lesser	–	–	–	–
	Both	–	3	–	2
4	None	–	–	–	–
	Greater	–	11	–	4
	Lesser	–	–	–	–
	Both	–	1	–	–
5	None	–	–	–	–
	Greater	–	–	–	–
	Lesser	–	–	–	–
	Both	–	–	–	2
10	None	–	–	–	–
	Greater	–	2	–	–
	Lesser	–	–	–	–
	Both	–	2	–	–

a Totals for each category thus equal twice the number of fractures in that category.

The two fractures that were assigned to category 10 (anterior dislocation plus greater tuberosity involvement) were judged to involve both the surgical neck and the greater tuberosity by both raters. Raters also agreed that one of the fractures involved the lesser tuberosity. On the study eligibility forms both fractures were assessed to involve the greater tuberosity only.

Agreement between the baseline assessment of tuberosity involvement and the Neer classification

The agreed Neer classifications were compared against the radiography assessments at baseline on the study eligibility forms. *Table 75* tabulates high-level groupings of the two assessments [Neer one- and two-part fractures and Neer three- and four-part fractures against tuberosity involvement at baseline (yes/no)], used for subgroup and sensitivity analyses.

TABLE 75 Agreed Neer classification and tuberosity assessment at baseline: overview

Agreed Neer classification	Tuberosity involvement (study eligibility form), <i>n</i> (%)	
	None (<i>n</i> = 57)	Greater and/or lesser tuberosity (<i>n</i> = 193)
Neer one- and two-part fractures	53 (93.0)	93 (48.2)
Neer three- and four-part fractures	4 (7.0)	100 (51.8)

The majority of the baseline assessments indicating no tuberosity involvement (93%) were identified as Neer one-part and two-part fractures. Conversely, half (48%) of the patients with an estimated involvement of one or both tuberosities at baseline were associated with Neer one-part and two-part fractures and the other half (52%) were associated with three-part and four-part fractures. Accordingly, the agreement between the two assessments was only fair (61% agreement, kappa = 0.29, $p < 0.001$).

Table 76 provides further details on how the assessment of tuberosity 'involvement' at baseline and the Neer classification of 'displacement' related to each other. The focus here is on the main expected and prevalent categories: 3 ($n = 119$), 8 ($n = 90$) and 12 ($n = 11$). If we equate here the term 'involve' with 'displacement meeting the Neer criteria', then Neer category 3 fractures (two-part surgical neck fractures) are not expected to 'involve' either tuberosity (the tuberosities are not displaced relative to the other parts of the humerus). This was true for 43% (51/119) of category 3 fractures according to the baseline assessment, with most of these fractures also being assessed to involve either the greater tuberosity only at baseline (34%) or both tuberosities (16%). Neer category 8 fractures were expected to 'involve' the surgical neck and greater tuberosity. This was true for the majority (54/90; 60%) of category 8 cases, with many cases (36%) judged to additionally 'involve' the lesser tuberosity at baseline. Neer category 12 (four-part) fractures were expected to 'involve' both tuberosities. This was true for 45% (5/11) of category 12 cases at baseline, with four fractures (36%) judged to 'involve' the greater tuberosity only at baseline.

TABLE 76 Agreed Neer classification and tuberosity assessment at baseline: detail

Agreed Neer classification			Tuberosity involvement (baseline study eligibility form), n (%)			
			None ($n = 57$)	Greater tuberosity ($n = 119$)	Lesser tuberosity ($n = 10$)	Both ($n = 64$)
1	Neer one part	Undisplaced surgical neck	2 (3.5)	12 (10.1)	0 (0.0)	4 (6.3)
3	Neer two part	Surgical neck	51 (89.5)	41 (34.5)	8 (80.0)	19 (29.7)
4		Greater tuberosity	0 (0.0)	6 (5.0)	0 (0.0)	2 (3.1)
5		Lesser tuberosity	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.6)
8	Neer three part	Surgical neck + greater tuberosity	3 (5.3)	54 (45.4)	1 (10.0)	32 (50.0)
9		Surgical neck + lesser tuberosity	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.6)
10		Anterior dislocation + greater tuberosity	0 (0.0)	2 (1.7)	0 (0.0)	0 (0.0)
12	Neer four part	Surgical neck + greater tuberosity + lesser tuberosity	1 (1.8)	4 (3.4)	1 (10.0)	5 (7.8)

a Bold font indicates the expected categories as part of the trial.

Chapter 7 Economic evaluation

Objective

The objective of this economic evaluation was to assist decision-making in determining whether or not surgery represents a cost-effective alternative within the UK NHS for the treatment of adults with an acute closed displaced fracture of the proximal humerus with involvement of the surgical neck.

Overview

Economic evaluation supports decision-making when prioritising the allocation of limited health-care resources.⁷⁰ Economic evaluation alongside clinical trials, as in the ProFHER trial, can therefore be a valuable tool to help decide what interventions should be implemented based not only on effectiveness but also on cost-effectiveness. Moreover, RCTs are often the best means for providing unbiased estimates of both health effects and costs.⁷¹ The pragmatic multicentre design of the ProFHER trial has the advantage of reflecting actual practice; hence, the cost profile of the participants and the interventions is likely to mirror clinical practice in UK NHS hospitals.

Individual patient data collected alongside the ProFHER trial were used to perform a cost-utility analysis in which health-related quality of life (HRQoL) was measured in terms of QALYs, which represent years lived in perfect health. Differences in mean costs and mean QALYs at 2 years were used to derive an estimate of the cost-effectiveness of surgery and non-surgical treatment. Costs and QALYs were evaluated on the basis of costs falling on the NHS and Personal Social Services (NHS perspective) and were expressed in UK pounds sterling at 2012 prices. Costs and QALYs were discounted from year 1 using the recommended discounting rates (3.5%) according to the current available guidance.⁵⁵ The analysis was conducted on an ITT basis; thus, the treatment groups were compared based on the initial random allocation of participants irrespective of protocol deviations or withdrawals.

Because of the magnitude of missing data throughout the trial follow-up, the base-case analysis was conducted using the multiple imputed data set and a sensitivity analysis of complete cases was carried out to test the impact of excluding patients with missing data on the final results.

Methods

Cost data sources

The data required for the analysis, both costs and health benefits (measured in terms of QALYs), were collected from individual patients and through health professionals during the 2-year follow-up period of the ProFHER trial.

Data on utilities were estimated by means of the EQ-5D questionnaire data elicited from patients at baseline (see *Appendix 17*) and 3 (see *Appendix 18*), 6, 12 and 24 months (see *Appendix 19*).

There are two main cost components in the analysis: (1) the cost of the surgical intervention and (2) the cost of the health-care consultations at baseline and at the 3-, 6-, 12- and 24-month follow-ups:

- Data extracted from the surgical forms (see *Appendix 21*), which include the main items of resource use relating to each operation, were used to calculate the cost of the surgical intervention for participants in the surgery group and that for six participants in the non-surgery group for whom surgical forms had been completed.

- The cost of the health-care consultation consists of all of the costs of visits to both primary and secondary health-care professionals. Resource use from primary health-care consultations was collected using patient questionnaires only. However, hospital outpatient visits, hospitalisations and physiotherapy appointments were recorded by two sources (patient questionnaires and hospital forms). Although physiotherapy treatment logs (see *Appendix 23*), completed by physiotherapists providing patient care, recorded information for only the first episode of care, these were a more complete record of physiotherapy that included data on the timing of individual sessions. We used these as the primary source for physiotherapy data. The other hospital data sources were the inpatient episode form (see *Appendix 22*) and the 1-year and 2-year follow-up forms (see *Appendices 25* and *26*). The inpatient episode form covered hospitalisation from initial surgery to discharge and hence non-surgical patients had an inpatient episode form only if they were admitted to hospital. The two hospital follow-up forms recorded any hospitalisation from discharge after initial treatment to the end of the second year because of surgical or medical complications. Both patients and health professionals were asked to determine whether visits and hospitalisations reported were shoulder or non-shoulder related.

Our economic analysis plan indicated that, when data could be sourced from patient questionnaires and hospital forms, hospital forms would be used as the main source for calculating resource use. This initial decision was made because the hospital forms provided more detailed data on specific incidents and these were completed by health-care professionals trained in the study and who were being paid for the return of completed forms. This decision was justified by the extent of missing data on resource use in patients' questionnaires.

Health-related quality of life and quality-adjusted life-years

Health-related quality of life was expressed in terms of utilities, which were obtained from trial patients using EQ-5D scores at baseline, 3 months, 6 months, 1 year and 2 years.

The EQ-5D⁷² is a standardised and validated generic instrument for the measurement of HRQoL that allows the translation of patient utilities into QALYs. As well as being one of the most used generic health status measures, it is the instrument recommended in National Institute for Health and Care Excellence (NICE) technology appraisal guidance.⁵⁵ This instrument considers health (functioning) in terms of five dimensions: mobility, ability to self-care, ability to undertake usual activities, pain and discomfort, and anxiety and depression. Each dimension has three possible levels: no problems, moderate problems or severe problems. This five-domain and three-level system generates 245 mutually exclusive health states, including unconscious and death. According to the responses to the EQ-5D classification system, a health status can be defined and a single index utility assigned. Each of these health states has been validated in a large UK population sample using the time trade-off method, ranging from 1 for perfect health (thus the maximum value possible) to -0.594 for severe problems; 0 corresponds to death.⁷³

The validity of the EQ-5D has been evaluated within the context of a prospective study that tested its ability to capture clinically important changes in patients with proximal humeral fractures.⁷⁴ This study found that the EQ-5D displayed good internal and external responsiveness and recommended its use as a quality of life measure in patients with these injuries.

We converted the utilities derived from the EQ-5D into QALYs for each patient using the AUC method, following the trapezium rule, which assumes linear interpolation between follow-up points.⁷⁵ Incremental mean QALYs between treatment groups were estimated with regression models according to treatment allocation. Despite the randomisation process, which ensures that baseline variables are balanced between the arms of the trial, in practice (regardless of sample size) it is normal to find an imbalance in mean baseline utility. As baseline utility is likely to be correlated with QALYs gained over time, there are robust reasons to control for baseline utility when estimating QALYs.⁷⁶ Therefore, we conducted two types of analysis: (1) based on adjusted baseline utility scores and (2) adjusted for a number of covariates. We used the same set of covariates as in the clinical effectiveness analysis, that is, age, gender, treatment group and tuberosity involvement (yes/no) at baseline.

Health-care resource use and costs

Surgery: data collection and cost estimates

On designing the surgical form (see *Appendix 21*) we were mindful of its key role in collecting information on resource use relating to surgery. The form collected information on operation times, the grade and number of staff involved in the operation, the type of implant used, the disposables required and whether or not there were any unexpected procedures during the intervention.

We received surgical forms for the 109 patients allocated to surgery who received surgery and six patients allocated to non-surgery who received surgery. Two of these patients were considered crossovers; the other four patients underwent surgery following unsuccessful conservative treatment.

Two key areas of missing data in the surgical forms for which data were imputed were operation time and number of staff involved in the operation. After discussion it was agreed that the mean operation time calculated from the rest of the data set should be used for the six patients with missing operation time. As mentioned in *Chapter 3*, data entry for numbers of personnel was problematic; frequently staff ticked the relevant boxes instead of entering the number of staff. In these cases we assumed that there was one surgeon or one anaesthetist in the theatre, as this was the typical (average) number of staff when provided.

Unit costs for surgery, together with their sources, are provided in *Table 77*. Aside from for implants, the unit costs used in the analysis were obtained from published national averaged tariffs: *Unit Costs of Health and Social Care* [Personal Social Services Research Unit (PSSRU)],⁷⁷ *NHS Reference Costs*⁷⁸ and the *British National Formulary* (BNF).⁷⁹

The staff cost per minute was estimated using PSSRU data⁷⁷ for each of the staff categories involved in the surgical intervention. For those categories for which direct information about the cost per hour was not provided, we used mean total earnings and total working times to infer the cost per minute. For example, the mean total salary per full-time equivalent for a senior house specialist is £41,100; based on a working time of 42.8 weeks per annum and 37.5 hours per week, we estimated a unit cost of £0.43 per minute for time spent in theatre. Radiographer time was costed for a patient's time in theatre only when an image intensifier was used.

Several of the hospitals involved in the ProFHER trial provided the prices (2012–13) that the hospital paid for the implants used. These costs excluded value-added tax (VAT) and included any discount that might have been agreed locally. Five different types of plates and screws provided by different manufacturers were used in the trial: PHILOS, AxSOS, S3 plate, Polarus PHP and NCB plate. Costs were obtained from the hospitals for three of the four hemiarthroplasties used – Epoca, Anatomical and Global FX/Advantage – and for the two nailing systems, which were used in just four patients – Polarus and Expert. Screws used in plates are costed individually. We based our cost estimate on an average of nine screws per plate, six in the humeral head fragment and three in the shaft. When the number of holes (three or five) was not reported for a plate, we either assumed that the plate used was the most frequently used plate type in that centre or, when this information was unavailable, used the average price from all hospitals to cost the plate. The prices for the implants detailed in *Table 77* are based on the hospitals' average cost per implant category. Hospitals provided detailed information about the specific costs involved in all eight (one in the non-surgery group) unexpected surgical procedures, which assisted with costing.

We derived the costs of the drugs used in the operation based on the information provided by the lead centre in the trial (James Cook University Hospital, Middlesbrough); we do not anticipate significant deviations from these costs in other centres. The drugs used were propofol (200 mg), fentanyl (100 µg) and morphine (10 mg) for induction; the antiemetics ondasetron (4 mg) and dexamethasone (8 mg), the muscle relaxant atracurium (50 mg) and the reversal agents neostigmine (2.5 mg) and glycopyrrolate (2.5 mg). Costs also included those of two doses of antibiotics, typically a combination of cefuroxime (1.5 g) and co-amoxiclav (1 g).

TABLE 77 Unit costs (and sources) used to estimate the total cost for each individual patient: primary surgery

Resource use	Unit cost (£)	Source
Surgeon/anaesthetist consultant	2.86 (per minute)	PSSRU ⁷⁷
Surgeon/anaesthetist registrar	1.43 (per minute)	PSSRU ⁷⁷
Hospital radiographer	0.60 (per minute)	PSSRU ⁷⁷
Nurse band 7	1 (per minute)	PSSRU ⁷⁷
Nurse band 6	0.80 (per minute)	PSSRU ⁷⁷
Nurse band 5	0.70 (per minute)	PSSRU ⁷⁷
Nurse band 4	0.22 (per minute)	PSSRU ⁷⁷
Nurse band 3	0.20 (per minute)	PSSRU ⁷⁷
Nurse band 2	0.17 (per minute)	PSSRU ⁷⁷
Senior house officer	0.43 (per minute)	PSSRU ⁷⁷
Associate specialist	0.94 (per minute)	PSSRU ⁷⁷
Staff grade	0.74 (per minute)	PSSRU ⁷⁷
Surgical assistant B5	0.50 (per minute)	PSSRU ⁷⁷
Surgical assistant B6	0.64 (per minute)	PSSRU ⁷⁷
Surgical ward per night	301	Department of Health ⁷⁸
General ward per night	261	Department of Health ⁷⁸
Plate and screws three hole	444.28	Manufacturer price 2012 ^a
Plate and screws five hole	455.50	Manufacturer price 2012 ^a
Lock screw	59.8	Manufacturer price 2012 ^a
Cortical screw	17.14	Manufacturer price 2012 ^a
Hemiarthroplasty	904.23	Manufacturer price 2012 ^a
Nail	482.43	Manufacturer price 2012 ^a
Propofol 200 mg	4.18	BNF ⁷⁹
Fentanyl 100 µmg	0.60	BNF ⁷⁹
Morphine 10 mg	15	BNF ⁷⁹
Ondansetron 4 mg	1	BNF ⁷⁹
Dexamethasone 8 mg	2.80	BNF ⁷⁹
Atracurium 50 mg	6	BNF ⁷⁹
Neostigmine 2.5 mg	0.50	BNF ⁷⁹
Glycopyrrolate 2.5 mg	0.91	BNF ⁷⁹
Cefuroxime 1.5 g	5.05	BNF ⁷⁹
Co-amoxiclav 1 g	1.06	BNF ⁷⁹

^a All manufacturer prices were provided by hospitals.

The cost of surgery therefore included the costs of the surgical team, implants, consumables and unexpected surgical procedures and inpatient stay (data from inpatient episode forms).

Primary care, secondary care and physiotherapy: data collection and cost estimates

Numerical information on primary care (visits to primary care professionals: GP, practice or community nurse, occupational therapist or physiotherapist) and secondary care (hospital visits: as an outpatient, as an inpatient or day admissions) was collected using the patient questionnaires at 3 months (see *Appendix 18*), 6 months, 1 year and 2 years. As well as being asked to record the number of times that they stayed in hospital as an inpatient, patients were asked to record the total number of nights spent in hospital over all visits. We assumed that, consistent with the format of the questionnaire, missing answers (boxes left blank) to the second question when patients reported no hospital stay indicated no use of services, thus no overnight stays.

As already mentioned, the number of hospital visits and complications reported on the hospital forms at 1 and 2 years were used as the primary sources of resource use. Similarly for physiotherapy, the primary analysis is based on the number of sessions reported by physiotherapists rather than the number of visits recorded by patients. *Table 78* details the unit costs for the aspects of patient care covered in this section.

TABLE 78 Unit costs (and sources) used to estimate total cost for each individual patient: primary care, secondary care and physiotherapy

Resource use	Unit cost (£)	Source
Primary care		
Visit to GP	40	PSSRU ⁷⁷
Visit to practice nurse	11	PSSRU ⁷⁷
Visit to community nurse	13	PSSRU ⁷⁷
Visit to occupational therapist	44	PSSRU ⁷⁷
Secondary care		
Shoulder hospital stay	3550	Department of Health ^{78a}
Shoulder excess hospital stay	301	Department of Health ^{78b}
Non-shoulder hospital stay	2724	Department of Health ^{78c}
Non-shoulder excess hospital stay	261	Department of Health ^{78d}
Outpatient visit	106	Department of Health ^{78e}
Day case	681	Department of Health ^{78e}
Physiotherapist		
Physiotherapy session	34	PSSRU ^{77f}

a Averaged (elective and non-elective), weighted by activity levels across all trusts, using the relevant shoulder Healthcare Resource Group (HRG) codes selected for the analysis: HA61B, HA61C, HA62Z, HA63Z, HB61B, HB61C, HB62B, HB62C, HB63Z.

b Excess bed-day averaged (elective and non-elective) per activity across all trusts using the relevant shoulder HRG codes selected for the analysis: HA61B, HA61C, HA62Z, HA63Z, HB61B, HB61C, HB62B, HB62C, HB63Z.

c Averaged (elective and non-elective), weighted by activity levels across all trusts and specialties.

d Excess bed-day averaged (elective and non-elective) per activity across all trusts and specialties.

e Outpatient visits and day cases averaged per activity across all trusts.

f The average duration for a physiotherapy session in the ProHER trial was half an hour, whereas the unit cost provided by the PSSRU is for a 1-hour session. The base-case analysis is based on a 1-hour session. When this unit cost was halved in the model, as expected (as physiotherapy sessions are equal in both arms) it had no impact on the incremental results.

We calculated the cost for each patient in the trial by multiplying their use of health-care resources by the associated unit costs (see *Tables 77 and 78*). The total cost for the base-case analysis included only shoulder-related resource use. It comprised four main components: (1) visits to primary and community health-care professionals (GP, practice or community nurse and occupational therapist); (2) hospital visits (inpatient, outpatient and day cases); (3) physiotherapy sessions; and (4) the cost of the surgical intervention. Other scenarios were tested as part of the sensitivity analysis, in which we explored the impact of incorporating both types of resource use (shoulder and non-shoulder related) and using the patient questionnaires as the primary source of resource use data for the analysis.

Multiple imputation

Missing data is a common and serious problem in most economic evaluations associated with clinical trials. Not only missing forms but also incomplete forms reduce, often considerably, the quantity of data on resource use that are available for analysis. The problem is amplified when there are frequent assessments, as in the ProFHER trial, in which there were five assessments: at baseline and at 3, 6, 12 and 24 months. Unfortunately, the loss of just one cost component or EQ-5D index result for a patient means that the total costs or total QALYs for that patient are lost to the analysis.

Complete case assessment excludes all patients with any missing data; thus, only patients with an observed total cost and QALY data are included in the analysis. As well as the resulting sample usually being a much reduced sample of the original data, complete case analysis might be biased if the patients included in the analysis are not a random subset of all study participants.⁸⁰ An alternative method to address missing data in clinical trials is multiple imputation,⁸¹ which has been recommended as the appropriate method to reflect the uncertainty in the results of an economic evaluation because of missing data.⁵⁷

Multiple imputation resolves the missing data problem by substituting each missing value with a predicted value. The multiple imputation process follows three consecutive steps. First, the imputed data set is created, through the use of regression models to predict plausible values for the missing observations from the observed values. The process includes all of the covariates that might be associated with the 'missingness mechanism' (why the data are missing); these included sex, age, tuberosity involvement, centre, costs (surgery, physiotherapy, primary care, inpatient and outpatient visits) and utilities (at baseline and 3, 6, 12 and 24 months). Costs and utilities were simultaneously imputed in the model rather than imputed separately. Therefore, the registered covariates were used for both costs and utilities, with a regression model fitted for each variable with missing values, with the previous variables as covariates. Based on the resulting model, a new regression model is then estimated and used to impute the missing values for each variable. A random component is included to reflect the uncertainty around the predictions. Thus, multiple imputation reflects the uncertainty in the prediction of missing values while preserving the distribution and correlations in the data.⁸² These values are then used to fill in the gaps in the data set. This process is repeated for a finite m number of times (m being the number of imputations), creating m number of imputed data sets. In the second stage, each data set is analysed independently using complete case methods. Finally, the estimates obtained from each imputed data set are pooled together to generate mean estimates of costs and QALYs, variances and CIs using Rubin's rules, in such a way that the uncertainty around the predicted values is fully taken into account.⁸³ As there were missing data for both costs and EQ-5D scores, multiple imputation using chained equations was employed.⁸² This way, each variable is predicted with its own regression model. Each imputed data set is created by running the regression models over several cycles, in which each variable informs the prediction of the other variables.

The correct specification of the regression models is key to ensuring that the distribution of imputed values does not differ from that of the observed values, thereby providing unbiased estimates. The specification of the regression models depends on the type and distribution of the variable to be imputed. Costs and QALYs (the variables to be predicted for this analysis) are both continuous and not normally distributed. Two alternative methods are proposed to deal with this difficulty when using multiple imputation with

chained equations: data transformations and predictive mean matching.⁸² As no transformation was successful in the case of the ProFHER trial data, predictive mean matching was used. This method ensures that observed data are used to estimate a predictive model (using the specified covariates) but, instead of replacing missing values with the model-predicted values, the nearest observed value is used to fill the missing value. This guarantees that the imputed values are sampled from values in the original data set and therefore no imputed values will lie outside the bounds of the original data distribution.

Given the extent of missing data in the ProFHER trial we decided to use the multiple imputation data set, created using all available data and multiple imputation with chained equations, as the 'base case'. Meanwhile, the use of the complete case data set was explored in the sensitivity analyses.

Incremental analysis

This cost-effectiveness analysis aims to guide decision-making. Given that total health-care expenditure must be covered from a limited and fixed budget, the most informative estimate of cost for decision-makers is the mean cost per patient. Also relevant is the mean utility per patient. Therefore, the focus of this analysis was to estimate the difference in mean costs and mean QALYs between the two treatment groups in the trial.

We used a bivariate modelling approach for the analysis. The incremental mean utility and the incremental mean cost between the two treatments were estimated using seemingly unrelated regression equations for data on costs and QALYs. This bivariate method brings efficiency gains over unrelated ordinary least squares (OLS) regression⁸³ for three reasons: (1) it allows for explicit modelling of both costs and effects while allowing the inclusion of a set of different covariates in the two equations; (2) it exploits the existence of correlation between costs and effects; and (3) seemingly unrelated regression does not require a new regression for every value of the cost-effectiveness threshold.⁸⁴ Again, the same set of covariates as used in the clinical effectiveness analysis was used: treatment group, age, gender and tuberosity involvement at baseline. The baseline EQ-5D utility was also included in the utility regression to adjust for possible baseline imbalance.⁸⁵

The incremental cost-effectiveness ratio (ICER) was estimated as the difference in mean total costs divided by the difference in mean total QALYs from baseline to 2 years. According to standard cost-effectiveness decision rules four different eventualities are plausible when comparing incremental costs and QALYs.⁸⁶ If the new intervention provides a better outcome (positive incremental QALYs) at a lower cost (negative incremental costs) it is considered a dominant intervention and hence cost-effective. On the contrary, if the new intervention achieves a poorer outcome (negative incremental QALYs) at a higher cost (positive incremental costs) it is considered a dominated option and hence not cost-effective. Thus, the ICER is considered only if either intervention does not dominate, that is, both incremental costs and incremental QALYs are positive (or negative). In these last two situations, to determine whether or not the incremental health gain is worth the incremental cost the ICER needs to be compared against a threshold value. For positive incremental costs and QALYs (the most frequent situation in health technology assessment), an intervention will be considered cost-effective only if the ICER is lower than the threshold. According to NICE,⁵⁵ the willingness to pay (WTP) for an additional QALY ranges between £20,000 and £30,000. Therefore, if the results of this cost-utility analysis – the estimated cost per QALY – are below this threshold, surgery would be considered cost-effective and its use in the NHS would be recommended.

The ICER can be rearranged in terms of net benefit, a more intuitive way of expressing whether or not surgery's health benefits are worth the additional costs.⁸⁷ The net benefit can be estimated on the cost scale as the incremental health gain expressed in terms of cost minus the incremental cost of the intervention. The health benefits are translated into a monetary value using the cost-effectiveness threshold, that is, the incremental QALYs are multiplied by the WTP threshold. Therefore, the net monetary benefit (NMB) provides an estimate of the gain (or loss) in resources of investing in a particular intervention when those resources might be used elsewhere.⁸⁸ Current NICE guidance⁵⁵ recommends presenting the NMB using values of £20,000 and £30,000 per QALY for the WTP threshold. Thus, surgery would be considered cost-effective only if the NMB is positive.

Uncertainty

The uncertainty around the cost-effectiveness results is explored by means of sensitivity analyses. As already mentioned, the extent of missing data justified the use of the imputed data set as the base-case scenario. Nevertheless, the complete case was tested as part of a sensitivity analysis and the results presented and compared with the imputed results. Complete case analysis offers unbiased estimates only if the data are missing completely at random, that is, the probability of the data being missing is independent of both observed and unobserved values, whereas multiple imputation provides unbiased estimates only if the data are MAR, that is, the probability of the data being unobserved is dependent on the observed but independent of the unobserved values. If the missing completely at random assumption does not hold the results of the complete case analysis might be biased, but this would apply to the results of the multiple imputation only if the MAR assumption also did not hold. Thus, multiple imputation has the advantage over complete case analysis in that if data are MAR, multiple imputation will produce unbiased estimates.⁸¹

As stated, the base case was based on the imputed data set and included only shoulder-related visits and hospitalisations; moreover, hospitalisations were estimated using the information reported on the hospital forms. Two one-way sensitivity analyses were conducted to test the impact of (1) including all visits and hospitalisations regardless of them being classified as shoulder or non-shoulder related and (2) using patients' questionnaires as the main source for estimating hospital visits and overnight stays.

Finally, we used probabilistic sensitivity analysis to investigate the uncertainty associated with the mean difference in costs and QALYs between treatment groups using both the complete case and the multiple imputation data sets. Non-parametric bootstrapping⁸⁹ was used to plot the joint distribution of costs and effects (QALYs) on the cost-effectiveness plane and derive the cost-effectiveness acceptability curves (CEACs) to express the (Bayesian) probability that surgery is cost-effective as a function of the threshold WTP.⁹⁰ The bootstrap technique was used to sample with replacement from the original observed pairs of costs and effects, maintaining the correlation structure, to create a new data set with 5000 observations. For each bootstrapped resample, an estimate of the differential costs and QALYs was calculated. The 95% CIs for the differential estimates were estimated using bias-corrected non-parametric bootstrapping.

Results

Patient population

Although the proportions of complete assessments during the trial were relatively high in both groups (87% of patients in each group returned their questionnaires at 2 years), the number of patients with complete follow-up assessments for all periods was much lower. A total of 173 (69%) patients – 95 (76%) allocated surgery and 78 (62%) allocated not surgery – comprised the complete case for utilities, that is, information for all five EQ-5D dimensions was available for all five assessment points. A total of 118 (47%) patients – 58 (46%) allocated surgery and 60 (48%) allocated not surgery – had complete resource use information for the whole follow-up period. However, when considering both utilities and resource use, complete information was available only for 100 (40%) patients: 54 (43%) allocated surgery and 46 (37%) allocated not surgery.

Fourteen patients died during the trial period, nine (7.2%) in the surgical arm and five (4.0%) in the non-surgical arm. When there were missing data before their deaths, the multiple imputation process was applied in the same way as for the rest of the patients in the trial. The questionnaires that should have been received at the assessment points after their death were considered as part of the complete case, with zero resource use and zero utilities.

Health-related quality of life and quality-adjusted life-years

To estimate QALYs, the EQ-5D responses were converted into utilities and then translated into QALYs using the AUC approach. The EQ-5D is considered to be complete only if patients give a response to all five dimensions. *Table 79* shows the number of questionnaires returned by patients (including

TABLE 79 Health-related quality of life: numbers of questionnaires returned and completed EQ-5D scores and corresponding proportions of missing data by allocation and follow-up

Follow-up	Questionnaires returned, <i>n</i> (%)		Completed EQ-5D, <i>n</i> (%)		Missing EQ-5D (not returned + incomplete), <i>n</i> (%)	
	Surgery (<i>n</i> = 125)	Not surgery (<i>n</i> = 125)	Surgery (<i>n</i> = 125)	Not surgery (<i>n</i> = 125)	Surgery (<i>n</i> = 125)	Not surgery (<i>n</i> = 125)
Baseline	125 (100)	125 (100)	123 (98)	121 (97)	2 (2)	4 (3)
3 months	114 (91)	111 (89)	106 (85)	98 (78)	19 (15)	27 (22)
6 months	113 (90)	119 (95)	111 (89)	114 (91)	14 (11)	11 (9)
12 months	111 (89)	115 (92)	109 (87)	109 (87)	16 (13)	16 (13)
24 months	109 (87)	109 (87)	108 (86)	103 (82)	17 (14)	22 (18)

questionnaires with any dimensions missing) and the number with complete EQ-5D data at each assessment point. Generally, the number of questionnaires returned decreased with time, as did the number of complete cases. Exceptions to this trend were found at 6 months in the return of questionnaires in the non-surgery group and in the number of complete cases in both groups. Of note also is that there were several patients who did not return questionnaires at one follow-up point but who returned completed questionnaires at subsequent follow-ups. The extent of incomplete EQ-5D data because of missing data strengthened the justification for using the multiple imputation data set as the base case.

Table 80 summarises the mean EQ-5D scores reported at each follow-up point for all of the available cases. The analysis of utilities shows that patients in the not surgery group started from a lower baseline utility on average than those in the surgery group (0.38 vs. 0.43). However, at the end of the second year there was little difference in EQ-5D score between the treatment groups.

The overall distribution of EQ-5D scores (utilities) for the different follow-up points is illustrated in Figure 23. At baseline, the utilities ranged between –0.358 and 1.

The distribution of mean utilities over the 2 years for the two groups is shown in Figure 24. The surgery patients reported better HRQoL at baseline, at 3 months and notably at 6 months. However, the patients in the non-surgery group reported better HRQoL subsequently. All of the differences were small and the 95% CIs overlap at each assessment point.

Total mean QALYs were estimated based on individual patient's utilities. Table 81 summarises the mean QALYs and the difference between allocation arms for all available cases. At the end of the trial, patients allocated to not surgery obtained on average a higher QALY gain than patients allocated surgery.

TABLE 80 Health-related quality of life: EQ-5D scores at baseline and at follow-up points up to 2 years

Follow-up	EQ-5D mean (SD) score		Difference (surgery – not surgery) (95% CI)
	Surgery	Not surgery	
Baseline	0.43 (0.37)	0.38 (0.37)	0.053 (–0.39 to 0.14)
3 months	0.64 (0.25)	0.63 (0.24)	0.013 (–0.06 to 0.06)
6 months	0.69 (0.24)	0.63 (0.28)	0.038 (–0.03 to 0.10)
12 months	0.65 (0.30)	0.68 (0.28)	–0.034 (–0.11 to 0.04)
24 months	0.67 (0.30)	0.69 (0.31)	–0.022 (–0.10 to 0.05)

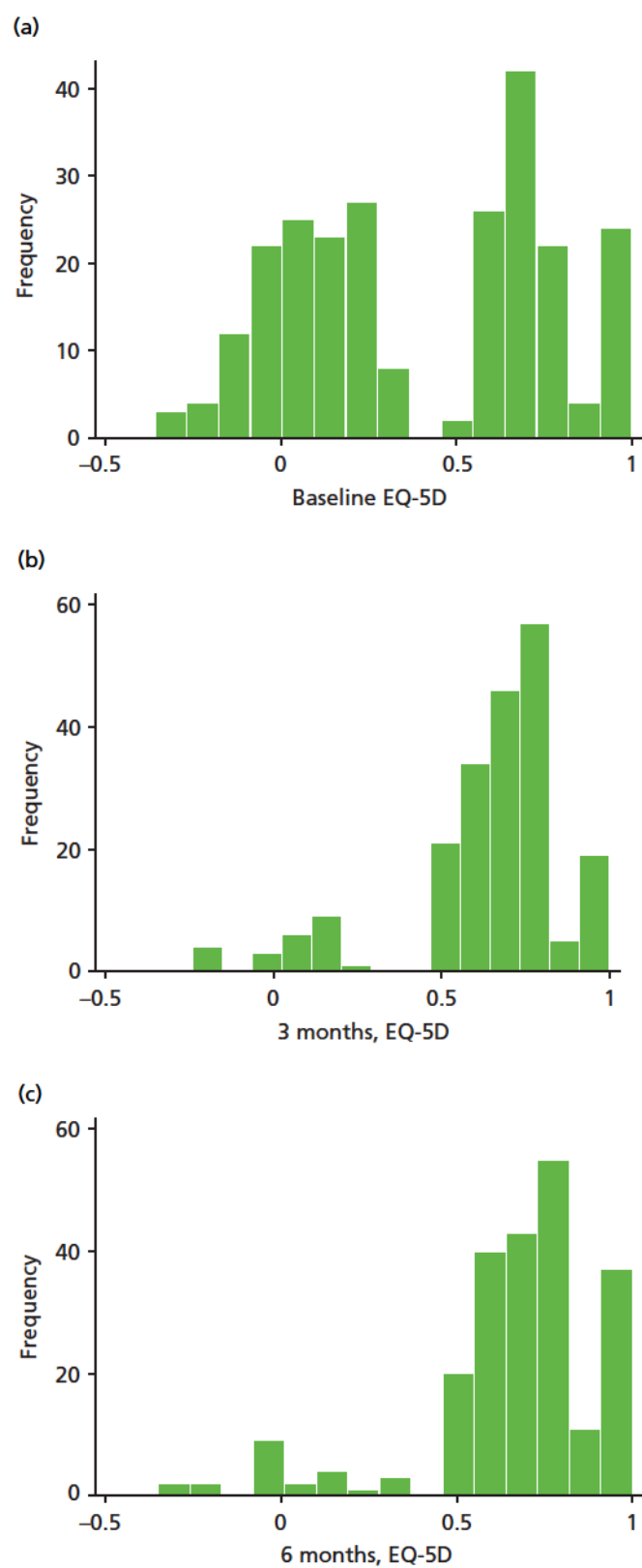


FIGURE 23 Distribution of EQ-5D scores at baseline and follow-up points up to 2 years. (continued)

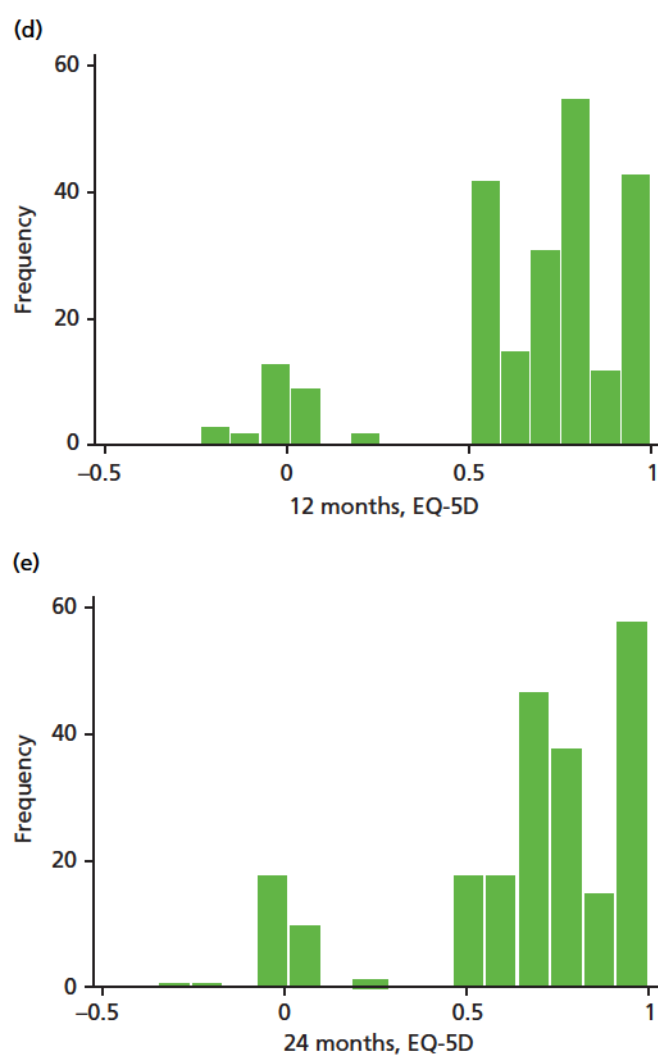


FIGURE 23 Distribution of EQ-5D scores at baseline and follow-up points up to 2 years.

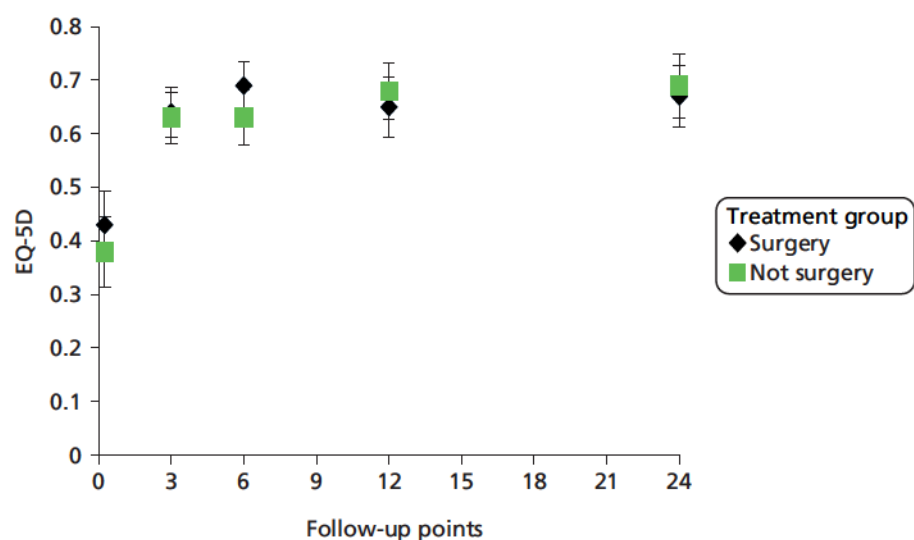


FIGURE 24 Mean EQ-5D scores at baseline and follow-up points up to 2 years. Error bars represent 95% CIs.

TABLE 81 Health-related quality of life: total QALYs for all available cases by allocation over the 2 years of follow-up

Allocation	Total	Mean (SD) QALYs	Min. QALYs	Max. QALYs	Difference ^a (surgery – not surgery) (95% CI)
Surgery	95	1.34 (0.43)	0.147	1.96	–0.066 (–0.186 to 0.054)
Not surgery	78	1.38 (0.37)	–0.029	1.94	

Max., maximum; min., minimum.
 a Estimated using OLS regression and controlling for baseline utility.

Health-care resource use and costs

The mean levels of resource usage over the 2 years of follow-up based on all available data are shown for the two treatment groups in *Table 82*. Although patients in the surgery group had on average fewer GP visits than non-surgery patients over the duration of the trial, they had on average more visits to the practice nurse, the community nurse and the occupational therapist. In terms of hospital resource use, patients in the surgery group had on average more outpatient appointments but fewer inpatient admissions than non-surgery group patients. The average number of inpatient days given in *Table 82* excludes the number of nights that surgery patients spent in hospital as a result of their initial surgical intervention, as this was included in the cost of surgery. Conversely, it does include hospital stay for non-surgery patients if they required hospitalisation. The average number of day cases is similar for both arms. Likewise, the number of physiotherapy sessions received did not differ between the treatment groups.

The resource use required for the surgical intervention was estimated in terms of the staff involved in the operation, the type of implant and disposables used, and the length of stay. Based on data from all 115 surgical forms received, the mean operation time in the ProFHER trial was 144 minutes, including anaesthesia time (114 minutes in theatre). The mean length of stay for surgery was 3.8 days.

The cost of surgery comprises the costs of the staff involved, the implants, the disposables (anaesthesia and antibiotics) and the inpatient stay. The nights that patients stayed in hospital as a result of the surgical intervention were costed at £301 per night using NHS Reference Costs for excess bed-days (average of the appropriate shoulder Healthcare Resource Group (HRG) codes for both elective and non-elective settings). The relevant HRG codes (relating to shoulder and upper arm procedures) in the NHS Reference Costs are HA61B, HA61C, HA62Z, HA63Z, HB61B, HB61C, HB62B, HB62C and HB63Z. The mean average cost of surgery in the trial was £3053 per patient (including elective and non-elective settings) for an average length of stay in hospital of 3.8 nights. The average cost for the selected HRG codes using NHS Reference Costs⁷⁸ is £2775 for the elective inpatient setting (average length of stay of 1.74 nights) and £3938 for the non-elective inpatient setting (average length of stay 4.14 nights). HES data¹⁰ related to surgery for a primary diagnosis of fracture of the upper end of the humerus performed in the UK during the last 5 years (from 2007/8 to 2011/12) show that the majority of patients are operated on as non-elective patients (one-third elective vs. two-thirds non-elective). From NHS Reference Costs,⁷⁸ the average unit cost for selected diagnosis codes weighted by activity levels and adjusted using the elective to non-elective ratio is £3550 (average length of stay 3.81 nights). Accordingly, a unit cost of £3550 was used in the analysis to cost the first 4 nights of the inpatient stay as a result of surgery to the shoulder and a unit cost of £301 was used for each of the remaining nights reported, that is, excess bed-days. To illustrate this point, consider a theoretical example of a patient who undergoes initial surgery requiring 3 days of hospitalisation and who then is admitted later for a complication related to the shoulder that requires 6 nights of hospitalisation. The cost of the inpatient stay for this specific patient would be the sum of the inpatient stay included in the cost of the surgical intervention – £903 (3 nights at £301 per night) – plus a later cost of the inpatient stay as a result of the complication – £4152 (4 nights costing £3550 plus 2 excess bed-days at £301 each). When interpreting the cost of a hospital stay as a result of surgery to the shoulder, it is important to appreciate that the hospital stay as a result of the primary surgical intervention is included

TABLE 82 Average resource use per arm of the trial over the 2 years of follow-up

Resource use	Surgery	Not surgery
Average number of GP visits		
<i>n</i>	76	71
Mean (SD)	0.85 (1.33)	1.18 (1.98)
Median (min., max.)	0 (0, 6)	1 (0, 12)
Missing, %	39	43
Average number of practice nurse visits		
<i>n</i>	74	75
Mean (SD)	0.78 (2.04)	0.28 (0.72)
Median (min., max.)	0 (0, 15)	0 (0, 4)
Missing, %	41	40
Average number of community nurse visits		
<i>n</i>	87	82
Mean (SD)	0.56 (2.31)	0.17 (1.33)
Median (min., max.)	0 (0, 16)	0 (0, 12)
Missing, %	30	34
Average number of occupational therapist visits		
<i>n</i>	86	81
Mean (SD)	0.66 (1.93)	0.59 (2.03)
Median (min., max.)	0 (0, 10)	0 (0, 12)
Missing, %	31	35
Average number of physiotherapy sessions		
<i>n</i>	118	117
Mean (SD)	9.57 (6.22)	9.60 (6.59)
Median (min., max.)	8 (1, 36)	8 (1, 43)
Missing, %	5	6
Average number of outpatient appointments		
<i>n</i>	106	112
Mean (SD)	0.41 (1.02)	0.34 (0.92)
Median (min., max.)	0 (0, 5)	0 (0, 5)
Missing, %	10	15
Average number of day case admissions		
<i>n</i>	114	114
Mean (SD)	0.096 (0.32)	0.096 (0.032)
Median (min., max.)	0 (0, 2)	0 (0, 6)
Missing, %	9	9
Average number of inpatient nights		
<i>n</i>	110	116
Mean (SD)	0.25 (1.23)	1.05 (3.15)
Median (min., max.)	0 (0, 16)	0 (0, 10)
Missing, %	12	7

Max., maximum; min., minimum.

within the cost of surgery and is thus not included in the total for the inpatient stay, which thus applies to subsequent shoulder complications. The same rationale was used to cost the number of nights in hospital as a result of non-shoulder-related or medical complications. From NHS Reference Costs,⁷⁸ the average unit cost across all inpatient settings is £2724 (average length of stay of 4 nights) and the average unit cost of excess bed-days across all trusts and specialties is £261. Therefore, the first 4 nights in hospital reported by patients as a result of medical complications were costed at £2724 and each subsequent night reported was costed at £261.

Table 83 summarises the mean cost by category of resource use and allocation based on all available cases. Costs associated with surgery were the major cost driver for the surgery group whereas the cost of hospital admissions was the cost driver for the non-surgical group.

Table 84 illustrates that patients in the surgical group experienced the largest proportion of the total cost during the first 3 months, that is, during the period when they received surgical treatment of their proximal humerus fracture.

TABLE 83 Cost by category and treatment allocation

Category	Mean (SD) cost (£)		Difference ^a (surgery – not surgery) (95% CI)
	Surgery	Not surgery	
Surgery	2565.85 (1633.64)	234.96 (1224.26)	2330.88 (1971.25 to 2690.52)
GP	34.21 (53.34)	47.32 (79.27)	–13.11 (–35.00 to 8.77)
Physiotherapy	325.56 (287.03)	326.63 (224.06)	–1.03 (–57.03 to 54.95)
Practice nurse	8.62 (22.46)	3.08 (7.99)	5.54 (0.09 to 10.98)
Community nurse	7.32 (24.80)	2.21 (17.3)	5.10 (–2.41 to 12.62)
Occupational therapist	24.78 (101.91)	7.51 (58.58)	17.26 (–8.17 to 42.71)
Hospital outpatient	43.40 (107.66)	36.04 (96.03)	7.41 (–19.79 to 34.61)
Hospital day case	65.37 (19.60)	65.37 (19.60)	0 (–54.61 to 54.61)
Hospital inpatient	341.25 (1198.36)	921.82 (2222.07)	–580.57 (–1052.18 to –108.96)

^a Estimated using OLS regression.

TABLE 84 Comparison of mean total costs per allocation during follow-up periods

Follow-up	Mean (SD) cost (£)		Difference ^a (surgery – not surgery) (95% CI)
	Surgery	Not surgery	
Month 3	2766.8 (1468.641)	694.00 (1868.78)	2072.79 (1595.84 to 2549.73)
Month 6 ^b	11.80 (32.81)	29.93 (140.70)	–18.13 (–46.78 to 10.52)
Month 12	183.42 (750.95)	231.21 (848.13)	–47.79 (–294.15 to 198.57)
Month 24	58.28 (366.95)	180.12 (708.44)	–121.83 (–288.72 to 45.05)
Physiotherapy	325.56 (287.03)	326.63 (224.06)	–1.03 (–57.03 to 54.95)

^a Estimated using OLS regression.
^b Resource use data at 6 months were collected only in the patient questionnaires and thus are exclusively related to primary care in the base-case analysis.

Cost-effectiveness analysis and uncertainty

The base-case analysis shows that the participants randomised to surgery accumulated greater costs and reported a lower HRQoL than participants randomised to non-surgery. *Table 85* shows the comparison of total mean costs and total mean QALYs between the multiple imputation results and the complete case data set. The similarity of both the mean and the CIs provides some reassurance of the validity of the multiple imputation model. Using multiple imputation, the surgical intervention cost on average £1780.73 (95% CI £1152.71 to £2408.75) more per patient than non-surgical treatment when adjusted for baseline utility. The incremental cost decreased very slightly when adjusted for the remaining covariates. Although total QALYs for the surgical intervention were smaller than those for non-surgery, the difference between treatment groups was not statistically significant either when adjusted for baseline utility (mean difference -0.0158, 95% CI -0.13 to 0.10) or when adjusted for covariates (mean difference -0.0101, 95% CI -0.13 to 0.11). The NMB associated with surgery was negative, indicating that this intervention is not cost-effective as the resources to be displaced would be greater than the benefit to be gained if surgery was implemented in the NHS.

As both the multiple imputation and complete case results indicate that surgery is dominated by the non-surgery intervention, this indicates that surgery should not be adopted. However, the differences in QALYs are not statistically significant, indicating that there might be some uncertainty associated with this conclusion. To analyse this uncertainty non-parametric bootstrapping was used. *Figure 25* shows the joint distribution of costs and effects for the 5000 bootstrapped replicates on the cost-effectiveness plane. The locations of the incremental cost pairs show that there is no uncertainty regarding the cost of surgery; surgery will always be a more costly intervention as all replicates fall above the horizontal axis. However, the position and spread of the QALY replicates indicate that there is uncertainty about the existence of a health benefit associated with surgery. This is consistent with the non-significant difference in QALYs between the two treatment groups. The outer ellipse in *Figure 26* represents the 95% CI for the true incremental cost-effectiveness for the compared alternatives. The inner ellipse defines the 50% confidence region. The point estimate represents the incremental cost (£1780.73) and incremental effect (-0.0158 QALYs) per patient based on the ProfHER data.

TABLE 85 Summary of the incremental analysis (ITT), cost-effectiveness results and uncertainty for the base case

	Multiple imputation		Complete case	
	Adjusted for baseline utility	Adjusted for covariates ^a	Adjusted for baseline utility	Adjusted for covariates ^a
Incremental mean cost (£) (95% CI) (surgery – non surgery)	1780.73 (1152.71 to 2408.75)	1757.65 (1126.29 to 2389.00)	1461.60 (527.10 to 2396.10)	1516.81 (614.94 to 2418.67)
Incremental mean QALYs (95% CI) (surgery – non surgery)	-0.0158 (-0.13 to 0.10)	-0.0101 (-0.13 to 0.11)	0.0004 (-0.16 to 0.16)	-0.0066 (-0.16 to 0.15)
ICER (£)	Dominated	Dominated	3,478,297.6	Dominated
NMB (£)	-2098.39	-1959.88	-1453.19	-1650.79
Probability cost-effective at WTP of £20,000 per QALY	0.05	0.06	0.20	0.16
Probability cost-effective at WTP of £30,000 per QALY	0.13	0.15	0.28	0.23

^a The covariates used to adjust for in the model were age, gender, treatment group and tuberosity involvement (yes/no) at baseline.

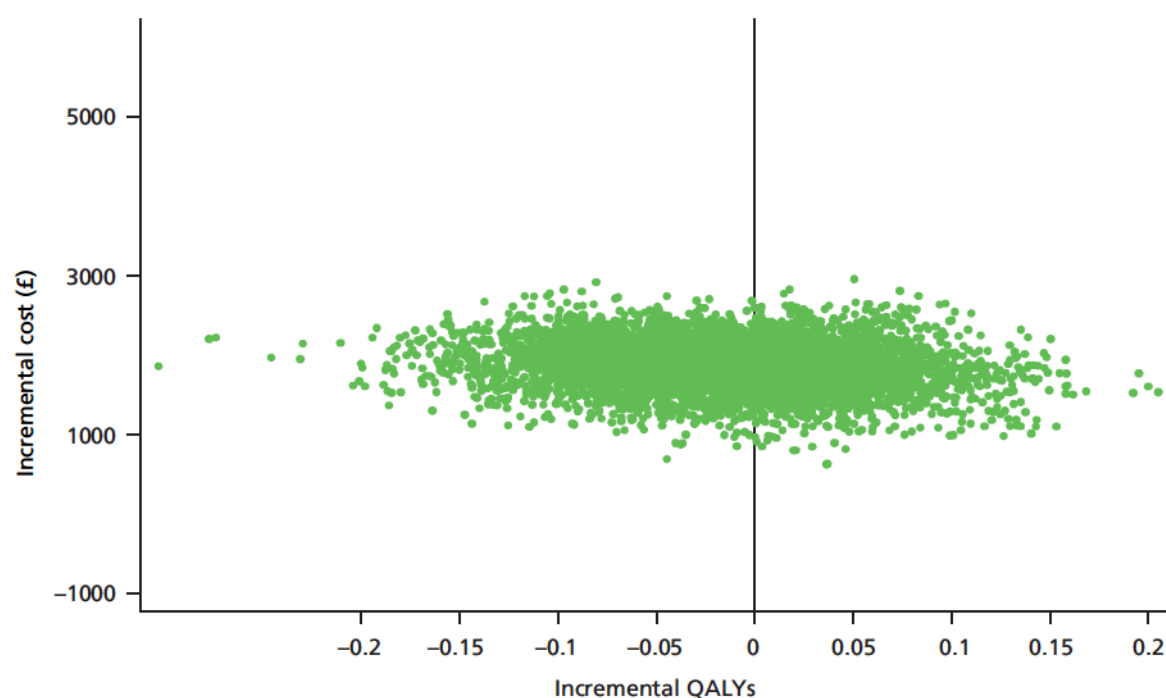


FIGURE 25 Cost-effectiveness plane, controlling for covariates: base-case analysis – non-parametric bootstrapping results.

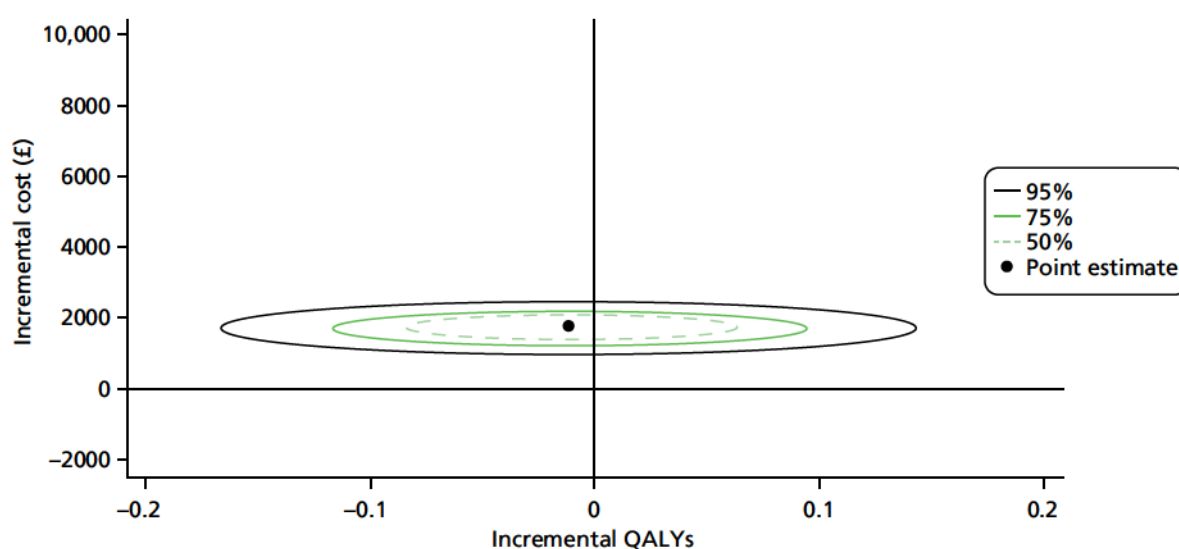


FIGURE 26 Confidence ellipses, controlling for covariates: base-case analysis.

The CEAC derived from the joint distribution of costs and effects is represented in *Figure 27*. The curve was constructed by plotting the proportion of the incremental cost-effectiveness pairs that are cost-effective for a range of thresholds. The y-axis, which presents the probability of surgery being cost-effective, ranges between 0% and 100%, with the horizontal dashed line indicating a 50% probability of surgery representing a value for money option for the NHS. As shown in *Figure 27*, the probability of surgery being cost-effective is < 10% (multiple imputation data set) given the current NICE WTP threshold of £20,000 per additional QALY. The probability is slightly higher for a threshold of £30,000 per additional QALY but is still too low for surgery to be considered as a cost-effective alternative by the NHS for fracture of the shoulder. Both of the adjusted analyses give similar results.

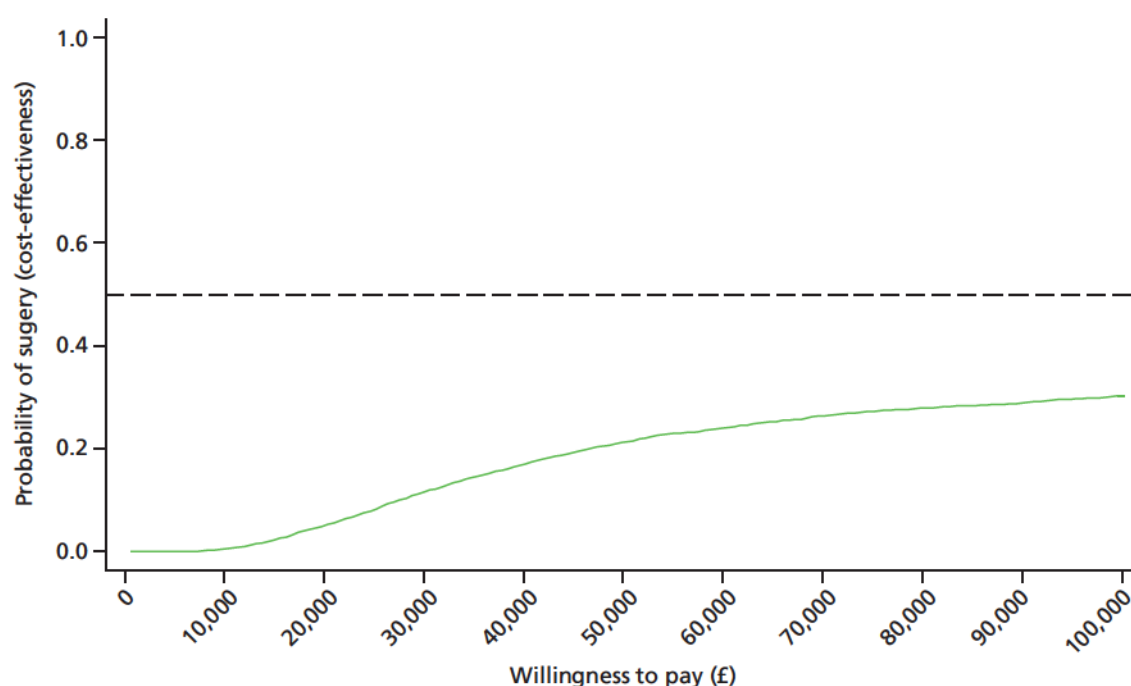


FIGURE 27 Cost-effectiveness acceptability curve, controlling for covariates: base-case analysis.

Sensitivity analysis

The complete-case analysis, whose results have been shown to be consistent with the results of the multiple imputation analysis, was tested as part of the sensitivity analysis.

The results of the one-way sensitivity analysis conducted to test the impact of including both shoulder- and non-shoulder-related resource use are shown in *Table 86*. The complete case for this scenario comprises 64 (26%) patients: 32 (26%) allocated to surgery and 32 (26%) allocated to non-surgery. There is little change in the results when considering all resource use in the assessment. Surgery is still not a cost-effective intervention for the multiple imputation data set. Although surgery does not represent a dominated option for the complete case, the ICERs for both adjusted scenarios are much higher than the thresholds that NICE normally consider for reimbursement decisions (£20,000–30,000 per QALY gained).

TABLE 86 Sensitivity analysis: summary of the incremental analysis (ITT), cost-effectiveness results and uncertainty when considering all resource use

	Multiple imputation		Complete case	
	Adjusted for baseline utility	Adjusted for covariates	Adjusted for baseline utility	Adjusted for covariates
Incremental mean cost (£) (95% CI) (surgery – non surgery)	1759.43 (906.30 to 2613.55)	1738.81 (908.63 to 2568.98)	909.95 (–1000.07 to 2819.98)	1312.38 (–606.22 to 3230.99)
Incremental mean QALYs (95% CI) (surgery – non surgery)	–0.01758 (–0.14 to 0.10)	–0.011 (–0.13 to 0.11)	0.0201 (–0.15 to 0.19)	0.0338 (–0.14 to 0.21)
ICER (£)	Dominated	Dominated	45,050.66	38,783.19
NMB (£)	–2111.19	–1970.41	–505.98	–635.60
Probability cost-effective at WTP of £20,000 per QALY	0.05	0.06	0.40	0.37
Probability cost-effective at WTP of £30,000 per QALY	0.13	0.15	0.42	0.41

The results are similar when we investigate the impact of using patient questionnaires (rather than hospital forms) as the main source for hospital data. The complete case for this scenario comprises 98 (40%) patients, 52 (42%) allocated to surgery and 46 (37%) to non-surgery. As shown in *Table 87*, there is no major impact on the results: surgery is always a dominated option, with a negative NMB and with a small probability of being cost-effective at the current thresholds.

Conclusion

Patient-level data from the ProFHER trial provide robust evidence on whether or not surgery is cost-effective in the treatment of adults with an acute closed displaced fracture of the proximal humerus with involvement of the surgical neck.

The base-case analysis (multiple imputation data set) for the ITT approach suggests that surgery is expected to be significantly more costly and to provide fewer health benefits than non-surgery for the treatment of these patients. Similarly, the analysis of uncertainty confirmed that it is unlikely that surgery represents an efficient intervention for the NHS, as the probability of surgery being cost-effective is ≤ 0.15 , regardless of the covariates used for adjustments. The results were robust to the three sensitivity analyses: (1) using the complete case data set; (2) considering patient questionnaires as the main source of hospital resource use; and (3) including both shoulder and non-shoulder-related resource use in the analysis. Although surgery was not a dominated alternative in the last scenario, the ICERs were still above NICE cost-effectiveness thresholds. In that sense, neither of these approaches (sensitivity analysis using patient questionnaires or sensitivity analysis using all resources), assessed for both the complete case and the multiple imputation data set, had any impact on the results.

TABLE 87 Sensitivity analysis: summary of the incremental analysis (ITT), cost-effectiveness results and uncertainty when considering patients' questionnaires

	Multiple imputation		Complete case	
	Adjusted for baseline utility	Adjusted for covariates	Adjusted for baseline utility	Adjusted for covariates
Incremental mean cost (£) (95% CI) (surgery–non surgery)	1609.44 (532.99 to 2685.88)	1563.10 (497.23 to 2628.98)	1740.65 (604.12 to 2877.19)	1792.93 (701.39 to 2884.47)
Incremental mean QALYs (95% CI) (surgery–non surgery)	–0.016 (–0.13 to 0.10)	–0.0103 (–0.13 to 0.11)	–0.0053 (–0.14 to 0.13)	–0.0120 (–0.15 to 0.12)
ICER (£)	Dominated	Dominated	Dominated	Dominated
NMB (£)	–1929.60	–1769.69	–1847.50	–2034.36
Probability cost-effective at WTP of £20,000 per QALY	0.08	0.10	0.12	0.09
Probability cost-effective at WTP of £30,000 per QALY	0.16	0.18	0.21	0.18

Chapter 8 Discussion and conclusion

The ProFHER trial is the largest RCT to date to examine treatment options for proximal humeral fracture, which is an increasingly common fracture in an ageing population. More specifically, it has effectively more than doubled the previous evidence available on whether or not surgery is required for the majority of displaced fractures of the proximal humerus. The ProFHER trial, with > 200 participants followed up at 2 years, is sufficiently powered to draw strong conclusions about the absence of a significant difference between the two groups in the primary outcome. Furthermore, using robust methods in the design of the trial with good implementation ensures strong internal validity. The pragmatic multicentre design of the ProFHER trial, including the constant emphasis on good standard practice, means that its results have immediate applicability in the UK NHS and are likely to apply to other countries with similar surgical practice.

In this discussion, we begin by summarising the main results, with guidance on their interpretation. We then explore whether or not there are potential risks of bias that might challenge the trial validity and applicability. In terms of the latter, we consider whether or not the trial population is representative of the putative population for which the treatment question applied and examine related issues of treatment preference and surgeon equipoise. The applicability of the findings to the NHS is discussed in the context of the findings of the trial and hospital episode data on the volume of operations at a hospital trust level. We provide a summary of the currently available evidence and pending RCTs examining this comparison for these patients. We conclude by considering the application of the trial results in practice, including how these findings can help the clinician make evidence-informed decisions in the clinical setting.

Summary of the main findings

Primary outcome

We found that there were no statistically significant differences in OSS [scale 0–48 (best outcome)] between the two treatment groups, either overall (difference of 0.75 score points in favour of the surgery group, 95% CI –1.33 to 2.84; $p = 0.479$) or at individual time points. Although not statistically significant, there was weak evidence to suggest a small difference in OSS at 6 months' follow-up (difference of 2.25 OSS points in favour of the surgery group, 95% CI –0.07 to 4.57, $p = 0.058$). However, the corresponding 95% CI excluded the prespecified difference of 5 points, suggesting that any difference is not of 'clinical significance'. The selection of the 5-point effect size is supported both for the general use of the OSS⁴¹ and, most recently, for its use in proximal humeral fractures.⁹¹ The latter study also reported an anchor-based (relating to patient perception of change) minimal clinically important difference of 11.4 points. Thus, both articles confirm that the difference between the two groups at 6 months is small and not clinically important.

As the OSS could not be used for baseline assessment, we cannot report on the responsiveness of the OSS in the period of greatest recovery. For example, a study of 20 patients with comparable characteristics to those of our trial participants found a change score of 15.6 (SD 10.7) for the OSS between 6 and 12 weeks.⁹¹ Nonetheless, the OSS improved over time in both treatment groups between 6 and 24 months, and checks of individual patients with low scores, indicating poor shoulder functioning, generally produced corroborative evidence of shoulder-related and other complications. Additionally, the small between-group divergence at 6 months was paralleled in the PCS domain of the SF-12 and the EQ-5D results. In the context of the greater number of newly treated medical complications occurring in the first year in the non-surgery group, it is plausible that the 6-month result is a reflection of greater morbidity, which impeded progress with shoulder functioning.

There was no statistically significant effect of treatment group for interactions with age or fracture type (assessed by tuberosity involvement at baseline and Neer classification) in the two planned subgroup analyses. Our prior expectations of directions of effect (subgroup differences) were not supported by these results but strengthen the case for not differentiating treatment (use of surgery) on the basis of these characteristics.

Similarly, the effect of treatment group remained not statistically significant when accounting for smoking status and patient treatment preference. The proportion of smokers recorded at baseline was greater in the non-surgery arm, putting these patients at greater risk of complications such as nonunion.⁹² However, this imbalance did not affect the results of the trial. Furthermore, a sensitivity analysis conducted after the preparation of this report showed that the effect of treatment group remained not statistically significant when accounting for clustering at centre level (see *Appendix 37*).

Almost half of the patients in the trial had no treatment preference at baseline. Interestingly, for these patients the difference between treatment groups in favour of surgery at 6 months' follow-up was the most pronounced, but still not statistically significant, compared with those patients who indicated a treatment preference for either surgery or not surgery. Although there is scope for further investigation of patient treatment preference, crucially this did not significantly affect patients' treatment outcome with regard to the main OSS findings.

Secondary outcomes

The use of a generic HRQoL measure such as the SF-12 in conjunction with the OSS is recommended practice.⁴¹ We found no evidence for statistically significant differences between treatment groups in physical functioning as measured by the PCS score of the SF-12 [scale 0–100 (best outcome); average 50]: the PCS score was on average 1.8 score points better in the surgical group than in the non-surgical group (95% CI –0.84 to 4.39; $p = 0.184$). The same lack of significance applied for mental functioning, as measured by the MCS score of the SF-12 [scale 0–100 (best outcome); average 50]: the MCS score was on average 1.3 points worse in the surgical group than in the non-surgical group (95% CI –3.80 to 1.23; $p = 0.317$). Although none of the differences between the two groups at individual time points were statistically significant for either of these outcomes, the results for the PCS mirrored those of the OSS analysis in that there was a slightly better score (mean difference 2.55) in the surgery group at 6 months. At most, this would equate to a small difference in favour of the surgery group but more notable is the similarity in the PCS scores over time in the two groups, with no additional improvement noted at 12 and 24 months for the surgery group.

Slightly more patients in the surgery group than in the non-surgery group experienced a surgical or shoulder fracture-related complication [30 (24%) vs. 23 (18%)]. Although the same number ($n = 11$; 9%) in each group had secondary surgery to the shoulder within the 2-year follow-up period, slightly more in the surgery group had increased or new shoulder-related therapy [seven (5.6%) vs. four (3.2%)]. None of these differences between groups (the number of patients with complications or the number undergoing further surgery or therapy) was statistically significant. The same finding applied to mortality, with there being slightly more deaths in the surgery group than in the non-surgery group [nine (7.2%) vs. five (4.0%)]. One death in the surgery group was judged as being trial related.

Although our collection of data on complications and their treatment was active and systematic, with specifically designed forms, the retrospective completion of forms at 1 year and 2 years could still be a source of missing data. However, although these outcomes may be under-reported, it is unlikely that this is a source of bias. This conjecture is supported by the findings of similar numbers in the two groups of additional 'untoward events' detected in our scrutiny of the physiotherapy logs and of prospectively collected adverse events, including those relating to the shoulder.

We recognise also that the named complications on the forms were mainly surgery related. However, the forms facilitated the collection of other 'complications', especially in relation to additional therapy or surgery. We have described our processing and interpretation of these complex data in *Chapter 3*, including, when possible, comments on the reported 'outcomes'. With the addition of other complications, notably post-traumatic stiffness (sometimes reported as 'frozen shoulder'), we believe that we have given an accurate account of the shoulder treatment and fracture-related complications incurred by these patients.

Cost-effectiveness

The base-case analysis showed that the participants randomised to surgery accumulated greater costs and reported a lower HRQoL than participants randomised to non-surgery. At 2 years, surgical intervention cost on average £1780.73 (95% CI £1152.71 to £2408.75) more per patient than non-surgical treatment (2012 prices). It was also slightly less beneficial, although this difference was not statistically significant [difference in utility of -0.015 (95% CI -0.13 to 0.10) when adjusted for baseline utility and -0.0101 (95% CI -0.13 to 0.11) when adjusted for covariates]. The NMB associated with surgery is negative, indicating that the resources to be displaced would be greater than the benefit to be gained if surgery was implemented in the NHS. Furthermore, the CEAC showed that surgery had only a 5% probability of costing < £20,000 to gain a QALY. Therefore, surgery appears to be a dominated treatment option and not a cost-effective use of health-care resources. These findings were robust to the three sensitivity analyses that were undertaken. Although surgery did not result in a dominated alternative for the base-case analysis or the sensitivity analyses that included both shoulder and non-shoulder-related resource use, the ICERs were still above NICE cost-effectiveness thresholds. This finding could have been influenced by an accurate account of the shoulder treatment and the greater number of newly treated medical complications occurring in the first year in the non-surgery group.

Countering potential threats to internal validity

We took all of the necessary measures to ensure trial validity and thus minimise the risk of bias. Some of these are discussed here.

Most importantly, these measures included a secure randomisation method, with a mid-recruitment decision not to proceed to minimisation to avoid any risk of predictability of treatment allocation. The comparability in the baseline characteristics of the two treatment groups attests to the success of this measure. A sensitivity analysis examining the chance imbalance in the numbers of smokers showed no effect on the primary outcome result.

We were proactive in establishing expectations for good standard clinical care, including clinicians using techniques that they were experienced with. We took steps to remedy identified gaps in the care programme for patients with these fractures, such as the absence of a patient information leaflet on sling immobilisation and a basic physiotherapy protocol. We stressed the need for comparability in the care provided to both treatment groups. Five key indicators of our success are listed here. First, at least two radiographic views were available for the majority of patients. Second, the majority of operations were conducted by consultant surgeons, with a consultant being in the theatre or available if needed for the rest of the operations, which were carried out by senior registrars. The choice of implants used, with the majority being locking plates (in particular PHILOS plates), is compatible with expectations of current practice. Third, the ProFHER trial sling information leaflet was provided, with very few exceptions, to all trial patients and to most of the patients who were eligible for the trial. Fourth, equivalent provision of physiotherapy was reported for the two treatment groups, as evidenced by comparability in the timing of sessions, the number of sessions and the interventions listed. Finally, physiotherapists recorded home exercises for the majority of patients for both treatment groups. Encouragingly, this showed that the majority of patients were performing their home exercises.

The high return rate for patient questionnaires and the almost complete return of hospital forms reflect the success of the measures taken to achieve high retention rates and as full a data set as possible. As described, particular effort, which was ultimately successful, was required to obtain portable copies of baseline radiographs. The baseline characteristics of patients contributing OSS data at 24 months showed no important differences between the two groups, with the exception of the continued difference between groups in the number of smokers. An analysis of the effect of missing data produced similar reassurance. Missing data from incomplete forms was a relatively minor problem in the clinical effectiveness analysis but, as explained in *Chapter 7*, the accumulative effects of missing data over several assessments meant that it was a serious problem in the economic analysis because of the strict criteria for complete case assessment, which excludes all patients with any missing data. The a priori decision to use multiple imputation as the base-case analysis to correct for missing data was thus correct.

An unavoidable limitation of trials such as the ProFHER trial that test sharply contrasting interventions is the lack of blinding of the assessment of patient-reported outcomes. Although this limitation is inescapable, measures to obtain as full a data set as possible and obtaining information on patient preferences help to assess the potential effect on the results of the lack of patient blinding. Altogether and separately, the similar and high rate of return of patient questionnaires at each follow-up point, the lack of differences between the two groups in baseline characteristics of those patients with OSS data at 24 months, the minimal effect of missing data on the analysis and the lack of a significant effect of baseline patient preferences on the OSS results give reassurance that the lack of blinding did not importantly affect patient responses overall. In addition, as described above, the difference between treatment groups in favour of surgery at 6 months was most pronounced in the 46% of patients who had no baseline treatment preference.

Rigorous methods were used for data processing and analysis. Both the clinical effectiveness analysis and the economic analysis were preceded by written documents that set out the planned analyses in detail. The former was overseen by the DMEC and, although the statistician was not blinded, the primary analyses were repeated by a second blinded statistician using different statistical software and the results were in agreement. ITT analysis, including all randomised patients in the groups to which they were randomised, was conducted throughout. There were more crossovers in the surgery group than in the non-surgery group [16 (13%) vs. two (1.6%)], but this is likely to reflect clinical practice, in which a patient can be found unsuitable for surgery subsequent to the fracture clinic consultation. Also, only a low number of patients were operated on by any individual participating surgeon; thus, no one surgeon dominated nor, indeed, any other clinician in providing treatment. Finally, coding, which always involved at least two independent raters, was carried out blind to allocation when this applied.

One commonly perceived threat to the validity of multicentre trials relates to the pattern of recruitment, including low recruitment from individual centres. Effective randomisation should counter any threat to internal validity, but the applicability of the results needs to be checked in terms of the characteristics of the actual trial population.

Applicability of the results

Characteristics of the trial population

The baseline characteristics of the trial participants and subsequent independent characterisation of radiographs confirm that the patients included in the ProFHER trial had sustained injuries that are typically considered for surgical intervention in contemporaneous practice. To emphasise this point a selection of four fractures that are representative of those in the ProFHER trial is shown in *Figure 28*. The results of this trial are therefore particularly relevant and important in guiding treatment decisions for these displaced fractures.

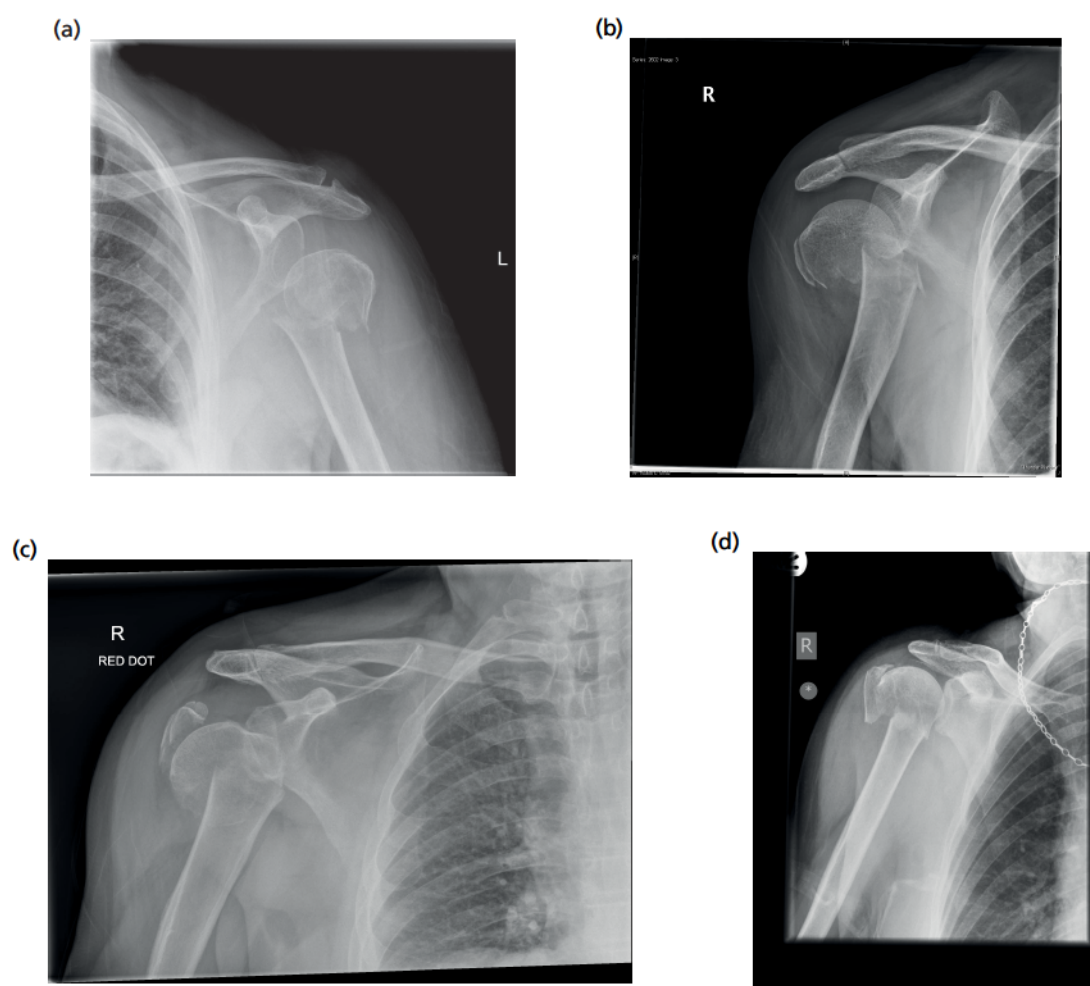


FIGURE 28 Four representative fractures from the ProFHER trial. (a) and (b) Two-part fracture of the surgical neck with an undisplaced greater tuberosity fracture; (c) and (d) three-part fracture with displacement of the surgical neck and greater tuberosity. L, left; R, right.

We aimed to include the majority of patients meeting the ProFHER trial's inclusion criteria but also established several exclusion criteria to avoid obstacles to appropriate (comorbidities precluding surgery or anaesthesia) or effective (not resident in catchment area) trial participation. A comparison of four key baseline characteristics revealed that patients who consented to participate in the trial tended to be on average 4 years younger (mean age 66 years) and more likely to have tuberosity involvement than those who were excluded. Although the higher age would also reflect exclusion because of comorbidities and lack of mental capacity, both of these differences were detected in the characteristics of the 117 patients excluded because of a lack of surgeon equipoise. However, although the surgeons tended to advise surgery in patients aged < 65 years, patients with fractures involving tuberosities were just as likely to be advised surgery or non-surgery as patients with fractures involving no tuberosities. This illustrates that these excluded fractures are still within the area of treatment uncertainty. We found no marked differences between enrolled and non-consenting patients. Overall, the age and gender characteristics of the screened population and the RCT population are comparable with those reported from other sources^{7,92,93} as well as HES data (see *Appendix 1*).

The results of the carefully prepared and performed independent classification of the fractures according to Neer confirmed that the study fractures all had 'involvement' of the surgical neck and that the majority met Neer displacement criteria. The formal classification fulfils the requirements for a description of the study fractures using the most commonly used classification system for these fractures. We accepted fractures that did not meet the Neer displacement criteria but which were nonetheless displaced fractures that would have reflected an individual surgeon's treatment uncertainty. The 18 one-part fractures, which

were distributed equally in the two intervention groups, were evidently displaced enough for the treating surgeon to have considered surgical intervention and hence they were included in the trial. The majority of these had greater tuberosity involvement. We consider that the inclusion of these fractures is not a threat to the internal or external validity of the trial. Ultimately, the Neer displacement criteria, although useful, are arbitrary and Neer⁵⁹ himself, in a comprehensive article reflecting on the 'four-segment classification', advised that 'As displacement is a continuum, there will always be some borderline lesions' (p. 392). It has been revealed also that Neer's criteria were devised in response to editorial instruction prior to publication of the original classification and were therefore additional to the consideration of the main study.²¹

Other than these 18 one-part fractures, the distribution of the fractures in the trial tended to be more towards the more complex fractures (i.e. three- and four-part fractures) than originally anticipated based on Court-Brown *et al.*'s epidemiological study.⁷ Based on data from this study, we would have expected 179 two-part fractures and 71 three- and four-part fractures (ratio of 2.52 : 1). The actual data showed 146 one- and two-part fractures and 104 three- and four-part fractures (ratio of 1.40 : 1). This, together with the other indications of severity (e.g. no contact surgical neck fractures) and the even smaller ratio of fractures with no tuberosity involvement (57 fractures) to fractures with any tuberosity involvement (193 fractures) as recorded on the study eligibility forms (0.30 : 1), points to the focus being on the more complex fractures. Although the distribution is shifted towards more complex fractures, it is in keeping with the aims of this pragmatic trial, which has addressed the treatment uncertainty that applied to the majority of displaced fractures of the proximal humerus. Thus, whether based on assessment of the fracture configuration (using radiography; CT scans were very rarely used) or a strict observation of Neer's displacement criteria, the study fractures are representative of the intended fracture population. Additionally, the absence of a significant effect on the study outcome in the subgroup analysis applied whether based on tuberosity involvement assessed at baseline or on the independent Neer classification. This indicates that the Neer classification is no more helpful than the assessment of tuberosity involvement/displacement; indeed, the latter, when exacting displacement criteria are not used, is more likely to be used as the basis for a surgeon's judgement in a busy clinic. Moreover, surgeons' inter-rater agreement on treatment has been found to be higher than the poor agreement found for the Neer classification.⁹⁴

Applicability to the UK NHS

The pragmatic design of the ProFHER trial lends itself for immediate application to the NHS setting. Our requirements, such as for surgeons to perform operations with which they were familiar and for there to be equivalent provision of physiotherapy, all came under the mantle of good standard practice in the NHS. We showed that consultant surgeons were invariably involved in the management of these fractures, and the majority of operations were carried out by these surgeons. Surgical fixation using locking plates and hemiarthroplasty, as used in the ProFHER trial, is the most commonly used method in current practice in the UK and in many other countries. We also showed that the majority of these patients received physiotherapy, with no distinction made between the treatment groups. The use of patient questionnaires and patient-reported outcome measures meant that no clinic follow-up assessment was required.

Thirty-five NHS centres, 33 of which are based in England, participated in the trial. An examination of centre variability by examining the OSS scores of patients in individual centres indicated greater variation within centres than between centres because of the low numbers of patients generally recruited. Although the small sample sizes for several individual centres hampers this sort of analysis, it gives support to considering the centres as a group. This we did in our investigation (based on using HES data for NHS hospitals in England) of the characteristics of the participating hospitals relative to those of non-participating acute hospitals in England. To obtain consistent data sets it was necessary to perform our analyses at hospital trust level, which is an overarching organisation providing particular NHS services in a particular region. We compared the general characteristics and those relating to surgical activity of patients with proximal humeral fracture in the 30 participating trusts (some centres belonged to the same trust) with those of the other 118–122 (year dependent) non-participating 'acute/teaching' trusts over a 5-year period. This covered the ProFHER trial recruitment period with an additional year at either end.

This analysis showed that a significantly greater proportion of participating trusts than non-participating trusts included teaching hospitals (0.33 vs. 0.14; $p < 0.0001$). Based on finished consultant episode (FCE) data, participating trusts in the ProFHER trial surgically treated, on average, significantly more patients with these fractures than surgeons in non-participating trusts [overall mean per year: 31.62 (SD 17.02) vs. 22.72 (SD 13.49); $p < 0.0001$]. These results are presented in *Figure 29*. The ratio of admissions to FCEs was nearly identical (0.90 vs. 0.89), which is reflected in the data showing a similarly higher level of admissions for episodes involving proximal humeral fracture surgery in the participating trusts (*Figure 30*).

These data show consistently higher surgical activity in relation to proximal humeral fractures in the participating centres throughout the 5-year period (39% overall). Comparative patient-level statistics are presented in *Table 88*. These show no important differences between the two groups of trusts in the age and gender of the patients undergoing surgery related to these fractures, nor in waiting time or length of stay. [Length of stay for both groups of trusts (HES data) was on average 2 days greater than that found for those treated surgically in the ProFHER trial (mean length of stay for surgery was 3.8 days).] The only difference was that participating centres were more likely to admit emergency patients than elective patients.

Overall, the HES data for surgically treated patients showed that these patients were on average 5 years younger than participants in the trial (mean 66.60 years). This is consistent with the general tendency to operate on younger patients, as also detected in surgeon treatment preference data on lack of equipoise and reflected in the increasing morbidity in older patients that would preclude surgery. Inspection of HES data for patients aged < 65 years and patients aged ≥ 65 years (as per our subgroup analyses) provides corroborative evidence. Summed over the 5-year period, in patients with a proximal humerus fracture who were aged < 65 years there were 9555 (43%) FCEs for operations out of the 22,084 FCEs for all possible treatments. For patients aged ≥ 65 years were 8821 (18.5%) FCEs for operations out of the 47,466 FCEs for all possible treatments.

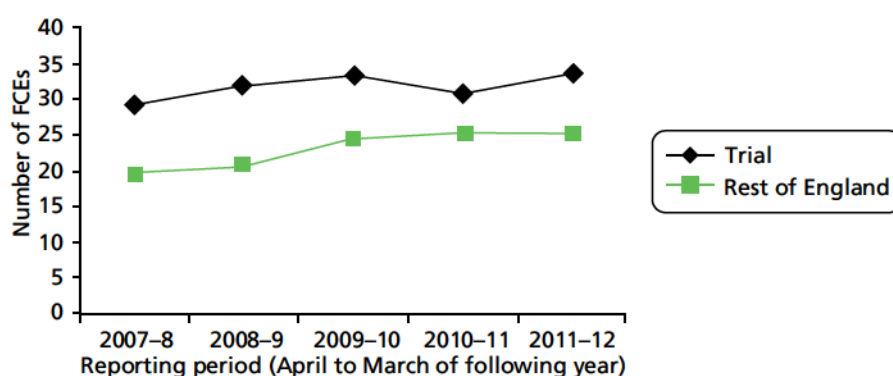


FIGURE 29 Average number of FCEs in participating and non-participating trusts.

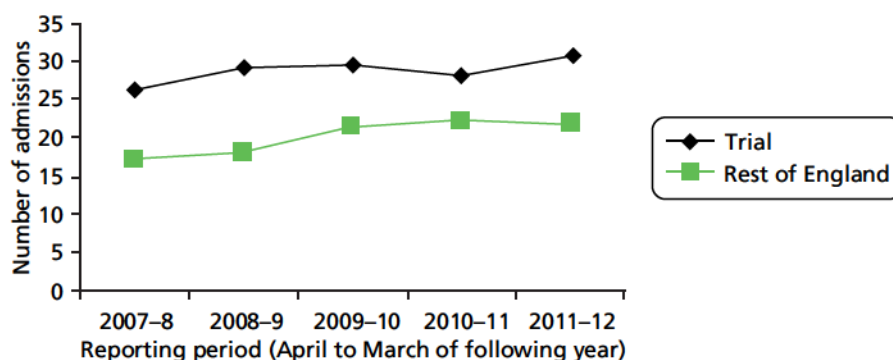


FIGURE 30 Average number of admissions in participating and non-participating trusts.

TABLE 88 Patient characteristics in participating and non-participating trusts

Variable	Non-participating trusts		Participating trusts		p-value
	Mean	SD	Mean	SD	
Age (years)	61.34	18.10	60.83	17.76	0.0923
Male ^a	0.34	0.47	0.34	0.48	0.6399
Emergency ^a	0.66	0.48	0.70	0.46	0.0000
Elective ^a	0.34	0.47	0.29	0.46	0.0000
Other ^a	0.01	0.09	0.01	0.10	0.1056
Day case ^a	0.01	0.08	0.01	0.09	0.0885
Waiting time (days)	7.73	19.86	7.67	15.22	0.9218
Length of stay (days)	5.73	8.91	6.00	8.59	0.0700

^a Proportion of whole sample.

Teaching hospitals and those with higher volumes of surgery are typically viewed as places that will feature more specialist surgeons who will perform more surgery. Hence, the results of the ProFHER trial apply not only to the participating centres but also to the rest of the acute hospitals in England and arguably to other countries with similar acute hospital settings and provision.

Subsequent to the ProFHER trial there has been a reconfiguration of trauma care in the UK, resulting in the establishment of major trauma centres. It is noteworthy that 11 of the 20 trauma centres for adults in England are located in hospitals that participated in the ProFHER trial.

Summary review of the currently available evidence

The evidence from other RCTs on this topic has been under regular review, also as part of a Cochrane review, and reports have been provided at TSC meetings. The currently published version of the Cochrane review³⁰ (search date January 2012) includes evidence from six RCTs (including 270 participants) comparing surgery with conservative (non-surgical) treatment plus details of six ongoing trials, including the ProFHER trial. One author (HH) performed an interim update of the evidence for this comparison. The update of the Cochrane review, which will commence on publication of this report, will follow the methods described in the review; this will include the independent reviewing of the ProFHER trial.

Using the search strategies listed in the Cochrane review, an interim search of the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE and EMBASE was conducted on 9 September 2013 by the Trial Search Co-ordinator of the Cochrane Bone, Joint and Muscle Trauma Group. HH also searched the World Health Organization (WHO) International Clinical Trials Registry Platform Search Portal (December 2012) and Current Controlled Trials (December 2012). HH screened a total of 219 records from the following databases: CENTRAL (36 records), MEDLINE (46 records), EMBASE (115 records), WHO Trials Registry (eight records registered from 2012) and Current Controlled Trials (14 records).

In total, HH identified one new single-centre RCT,⁹⁵ which compared hemiarthroplasty with non-surgical treatment in 50 patients, and a published protocol for an ongoing trial [Treatment of Proximal Humeral Fractures (TPHF)].⁹⁶ Brief details of the eight RCTs now available on this topic are provided in *Table 89*.

TABLE 89 Characteristics of the eight RCTs comparing surgery with conservative treatment for proximal humeral fractures

Study	Participants (Neer classification)	Surgery	Conservative (starting with)	Follow-up
ProFHER trial	250 patients with two-, three- or four-part fractures, all involving the surgical neck (UK)	Surgeons' choice according to expertise with implant: plate (PHILOS mostly), hemiarthroplasty, nails, other	Sling immobilisation	2 years
Boons 2012 ⁹⁵	50 patients with four-part fractures (the Netherlands)	Hemiarthroplasty (Global)	Shoulder immobiliser	1 year
Olerud 2011 ³²	60 patients with three-part fractures (all with surgical neck fracture) (Sweden)	Open reduction and fixation with a PHILOS plate and non-absorbable sutures	Sling immobilisation	2 years
Olerud 2011 ³³	55 participants with four-part fractures (Sweden)	Humeral head replacement with the Global FX prosthesis	Sling immobilisation	2 years
Fjalestad 2010 ³¹	50 patients with three- or four-part fractures (Norway)	Open reduction and fixation with an interlocking plate device and metal cerclages	Immobilisation of the injured arm in a modified Velpeau bandage. Closed reduction in eight patients	1 year so far
Zyto 1997 ²⁹	40 patients with three- or four-part fractures (three others excluded) (Sweden)	Internal fixation using a surgical tension band or cerclage wiring	Sling immobilisation	50 months
Kristiansen 1988 ²⁷	30 patients with 31 with two-, three- or four-part fractures (Denmark)	Percutaneous reduction and external fixation	Closed manipulation and sling immobilisation	2 years
Stableforth 1984 ²⁸	32 participants with four-part fractures (UK)	Hemiarthroplasty	Closed manipulation and sling	6 months

Results

Three outcomes are presented for the purposes of this report: patient-reported functional scores; additional or secondary surgery; and mortality. Pooled results for patient-reported functional scores at 6 months, 1 year and 2 years using Review Manager 5.2 (The Cochrane Collaboration, The Nordic Cochrane Centre, Copenhagen, Denmark) are presented in *Figure 31*. These results, which are dominated by those from the ProFHER trial, show no statistically significant differences between the two groups at any of the three follow-up times. Boons *et al.*⁹⁵ reported finding 'no between group differences' in Simple Shoulder Test scores at 12 months.

Pooled data from seven trials showed that significantly more surgical group patients than non-surgical group patients had subsequent shoulder surgery (30/262 vs. 16/261; risk ratio 1.82, 95% CI 1.03 to 3.22; $p = 0.04$) (*Figure 32*). This is equivalent to an extra operation in one of every 20 surgically treated patients. The equivalence between the two groups in the ProFHER trial in the incidence of subsequent shoulder surgery is unusual and may partly reflect a more inclusive approach to treatment and fracture-related complications. This is supported by evidence from a prospective 1-year study of 160 non-surgically treated patients, which reported that nine patients subsequently received surgery: four underwent surgical fixation and five had arthroscopic decompression.⁹² It is also clear that the number of surgically treated complications (11/109; 10%) in the ProFHER trial was not in excess. Further support for this conclusion comes from a systematic review of 12 cohort studies reporting results for 791 patients treated with locking plates for proximal fracture of the humerus; this found that 13.7% of these patients required a reoperation.⁹⁷

As shown in *Figure 33*, pooled data from six trials showed no significant difference in mortality between the two groups.

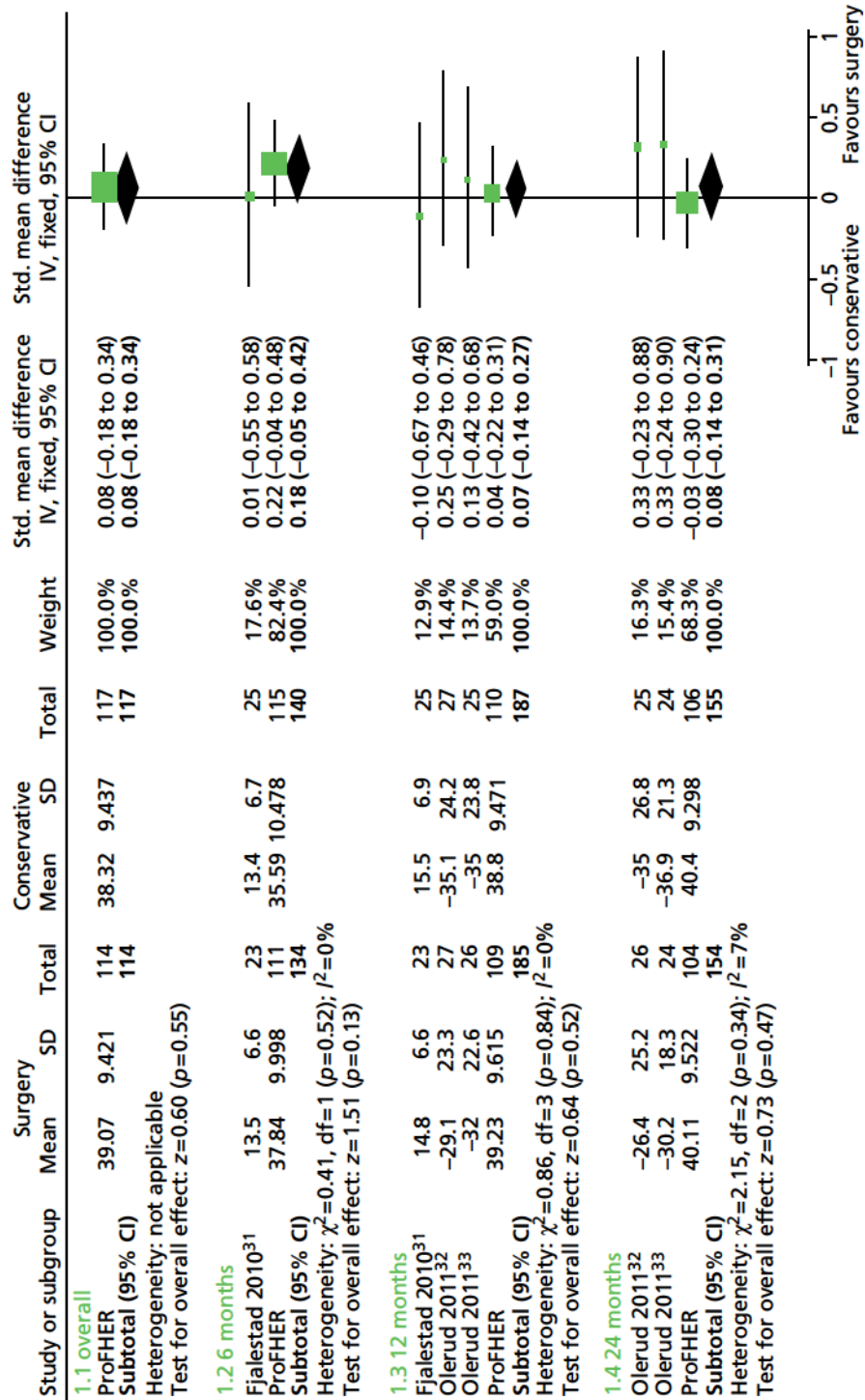


FIGURE 31 Meta-analysis of patient-reported functional scores. IV, inverse variance; std., standard.

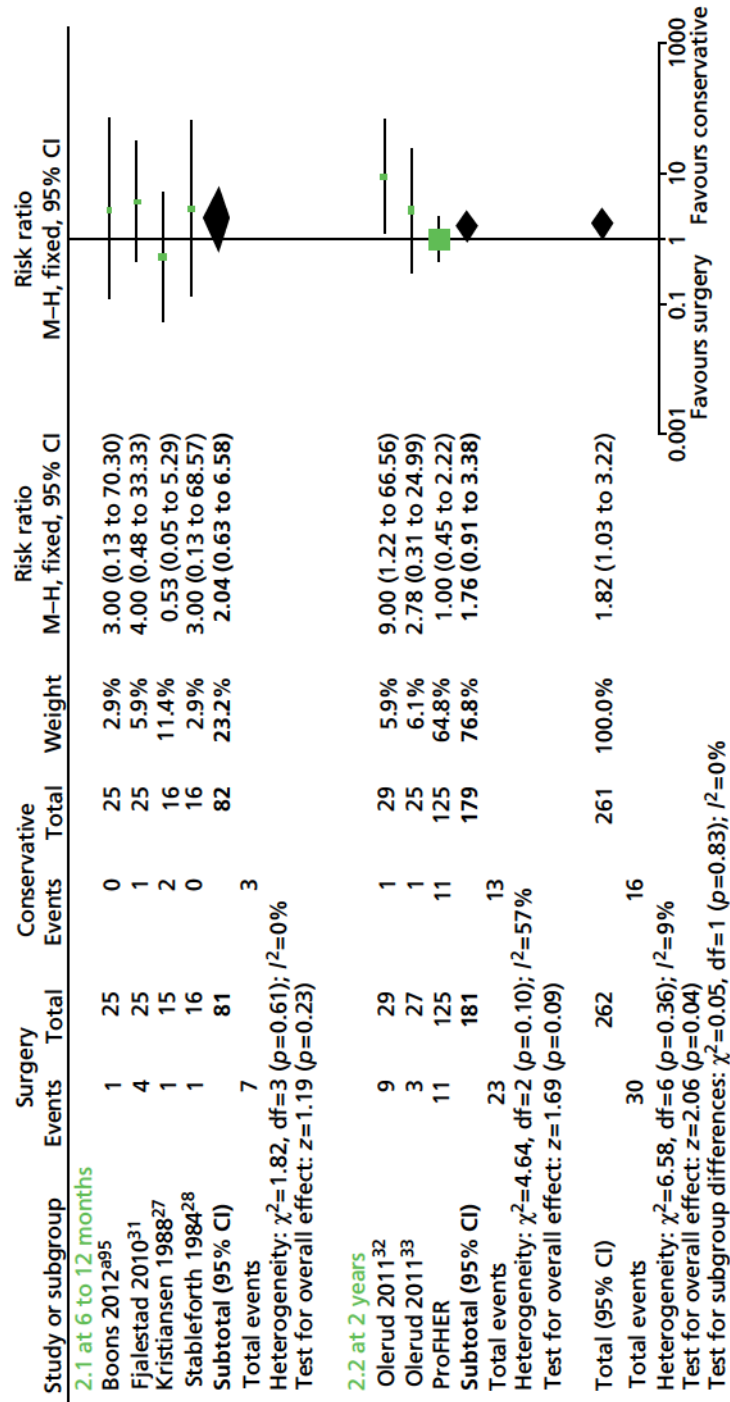


FIGURE 32 Meta-analysis of subsequent shoulder surgery. a, One patient in the conservative treatment group underwent an operation at 13 months for non-union. M-H, Mantel-Haenszel.

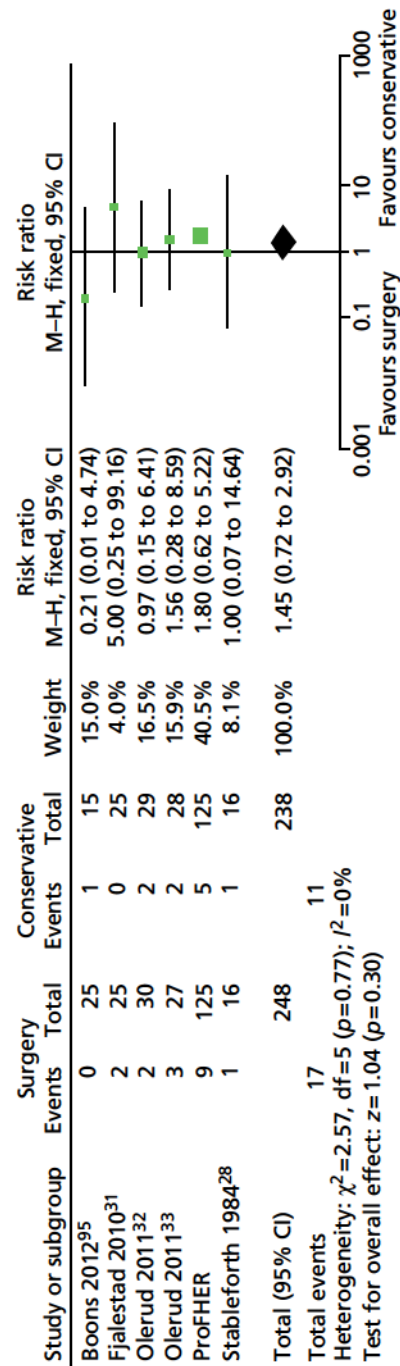


FIGURE 33 Meta-analysis of mortality. M-H, Mantel-Haenszel.

Pending and ongoing trials

A further five RCTs, based in four countries, are ongoing. Brief characteristics of these RCTs are presented in *Table 90*. A check carried out in December 2014 of the trial registration documents for these trials indicates that two are still recruiting whereas the recruitment status of three is unknown. These five RCTs plan to recruit a total of 802 patients aged ≥ 60 years, the majority with a three- or four-part fracture.

Application of the trial results to practice

In the above, we have provided justification and support from various sources that the results of our trial are reliable, are consistent with current practice and apply to the patient population represented by the trials inclusion and exclusion criteria. These criteria provide a preliminary guide to the patient population to which the trial findings apply. Thus, for this population, our trial has provided strong evidence that the majority of patients with displaced fractures involving the surgical neck, irrespective of age and involvement of tuberosity, are unlikely to benefit from surgery. Our results do not support the trend towards increased surgery in this group, which partly resulted from the availability of locking plates. Although we would not expect the trial results to completely alter surgical practice for these fractures, they should at least slow, stop or even reverse the increase in surgical intervention in the NHS. As well as preventing ineffective and thus unnecessary surgery, this would free up resources and operating time for fractures that should be treated surgically.

TABLE 90 Ongoing trials

Study	Location, start date ^a and recruitment status ^b	Participants	Interventions
Brorson 2009 ⁹⁸	Sweden (five centres) Start: April 2009 Recruiting status unknown	162 patients, age 60+ years, with four-part fractures	1. Hemiarthroplasty 2. Fixed-angle plate osteosynthesis 3. Non-surgical treatment
NCT00818987 ⁹⁹	Canada Start: November 2010 Recruiting status unknown	120 patients, age 70+ years, with three- or four-part fractures	1. Open reduction and internal fixation 2. Non-operative treatment (reduction and immobilisation)
NCT00999193 ¹⁰⁰	Finland Start: November 2010 Recruiting verified October 2014	150 patients, age 65+ years, with three- and four-part fractures	1. PHILOS locking plate 2. Hemiarthroplasty (Global FX) 3. Conservative treatment
Den Hartog 2010 ¹⁰¹	Netherlands Start: 15 June 2009 Recruiting status unknown	80 patients, age 65+ years, with a comminuted fracture (Hertel classification)	1. Hemiarthroplasty (Affinis®; Mathys Ltd, Bettlach, Switzerland) 2. Conservative treatment
Launonen 2012 ⁹⁶	Finland (four centres) Start: January 2011 Recruiting verified November 2014	290 patients, age 60+ years, with two-, three- and four-part fractures	1. PHILOS locking plate 2. Hemiarthroplasty (Epoca) 3. Conservative management

^a The start dates are the proposed start dates for patient recruitment at title registration.

^b This is based on information available in the trial registration documentation (checked 22 December 2014).

Crucially, this evidence can help to inform discussions with patients, allowing surgeons to reassure patients and perhaps themselves that surgery is not essential, thus avoiding some agonising decisions as well as unnecessary surgery. The majority of fractures in the ProFHER trial were complex but, although more complex fractures were recruited than anticipated, the number of confirmed four-part fractures was small ($n = 11$). These are rare fractures (2% of the population) and, although there is no evidence from the ProFHER trial or another recent trial⁹⁵ that the results would differ for these fractures, we acknowledge that there is uncertainty. The evidence from an ongoing multicentre trial⁹⁸ aimed at these fractures should help resolve this issue in due course.

As discussed, the strength of the pragmatic design of the ProFHER trial is that it represents current practice. However, as pointed out in *Chapter 1*, technology is changing all the time, with various pressures for early implementation. This raises the question of how applicable these results are or can be when the nature of surgical intervention has fundamentally changed. For instance, for the more complex fractures (including those outside the target population of the ProFHER trial) in older adults, there is a growing trend towards the use of reverse total shoulder arthroplasty, both in the UK and in other countries such as the USA.¹⁰² Reverse total shoulder arthroplasty, which was used once in the ProFHER trial as a revision procedure, is much more costly, more invasive and of unproven benefit for the patient population covered by the ProFHER trial. Until reliable evidence of its superiority is available, including in relation to hemiarthroplasty, the increasing use of the reverse prosthesis is questionable. Especially when there is a lack of evidence to inform the choice of surgical intervention, the pragmatic approach adopted by the ProFHER trial in which experienced surgeons perform operations with which they are familiar is a fundamental safeguard that underpins the continuing external validity of the trial results. The onus is on surgeons who suggest that these results no longer apply to provide the evidence that the new techniques are actually superior to the old.

Conclusion

The ProFHER trial has provided robust clinically relevant evidence showing that current surgical practice does not result in a better outcome for most patients with a displaced fracture of the proximal humerus involving the surgical neck. It is also not cost-effective in the UK setting.

It is important that non-surgical care should be of a good standard including the availability of a leaflet about sling immobilisation, timely access to physiotherapy and the promotion of home exercises. The potential need for remedial surgery for severe symptomatic complications in around 5–10% of these patients should be factored into forecasts for hospital budgets.

Recommendations for research

Given the findings of the ProFHER trial and the existence of five ongoing trials comparing surgical with non-surgical treatment, the initiation of further trials investigating this comparison is not warranted.

The growing number and importance of these fractures merits a more systematic and prospective collection of data on epidemiology, management and outcome; consideration is warranted as to the feasibility and importance of setting up a national database of these fractures, which should include the recording of patient-reported outcomes. This will help to remedy the serious lack of data on these fractures that hinders both practice and research in the UK. Although there are major differences between proximal humeral fractures and hip fractures in terms of the patient population, management, outcomes and the evidence base, the usefulness of such a database as a clinical audit tool that could facilitate improvements in the quality of care has been shown by the already established National Hip Fracture Database (NHFD).¹⁰³ Given the processes already in place for the NHFD, it may be more practical to extend the remit of this database to include proximal humeral fractures.

One important question for future research is, 'What are the best ways of informing patients with proximal humeral (and other upper limb) fractures about initial self-care?' People presenting to emergency departments are routinely sent home with their arm in a sling but, as we found out when conducting the ProFHER trial, no written information is provided. As described, we developed a trial-specific information leaflet on personal care during sling immobilisation. Research is required to establish the best approach to providing patient information on the early treatment of these fractures. Patients are often distressed by their injury and fail to ask questions or retain information; thus, the provision of written information seems appropriate but needs confirmation. The contents and format of such information needs input from patients; this also presents an opportunity to gather better information on current practice and obtain feedback from the patients on the practical considerations of (various types of) sling immobilisation.

Acknowledgements

Foremost, we thank the patients participating in this trial, without whom the trial would have not been possible. A general thanks also to all of those people whose commitment and efforts in the design and conduct of the ProFHER trial allowed us to successfully complete this study and this final report.

We are very grateful to the independent members of the TSC (Sarah Purdy, Andrew Carr and Brian Faragher) and the DMEC (Marion Campbell, Roger Francis and David Limb) for their resolute support and interest in the trial throughout and, in particular, their encouragement during the early challenges to site set-up and patient recruitment. We also thank the sponsor representative (Alasdair MacSween) and the patient representatives (Dorothy Anderson and Margaret Newlands) for attending the TSC meetings and for their enthusiastic support for the study.

To the staff at Teesside University, we thank Nigel Hanchard for his help, in particular with developing the sling immobilisation leaflet and the physiotherapy protocol. We are grateful to Lesley Watson for managing the contracts and research budget and to Paul Keane and Denis Martin for senior management support of the study at Teesside.

To the team at James Cook University Hospital (lead site), we are thankful to Lucy Micklewright, Birgit Hanusch and Lucksy Kottam for their significant contribution during planning, recruitment and data collection, which ensured the success of the study at the lead site.

To the team at the University of York, we are grateful to various staff who at different stages of the study contributed to the statistical (Martin Bland, Arthur Kangombe and Gill Worthy) and health economic (Ana Duarte, Jo Dumville and Gerry Richardson) components. This included contributions to the protocol (JD, GW), choosing the method of randomisation (MB), the design of the patient questionnaires (JD), a DMEC report (AK), checking the utility estimates of the EQ-5D (AD) and advising on the health economic methods (AD, JD, GR). We also thank the data managers (Ben Cross and Valerie Wadsworth) and the ProFHER trial secretaries (Sarah Gardner, Sue Collins, Sue Chapman and Angela Rogers) whose technical and administrative support, respectively, enabled the successful management of the study. We also thank Nils Gutacker for sharing his expertise and knowledge of HES data, which has informed several aspects of the final report.

We would also like to thank the clinicians who contributed to the study. This includes Brian Cox for his advice about collecting radiography images and measuring displacement; Martin Scott and Helen Watson for their contribution to developing the physiotherapy protocol; and Muhammed Ismail and Will Eardley for their assistance with piloting the training in the Neer classification of the radiographs. Special thanks are due to Jaime Candal-Couto, who fulfilled various roles in the study including being a principal investigator at a recruiting site, appraising baseline radiographs for quality and deputising for the chief investigator during his ABC fellowship; and Alan Johnstone and Andrew Brooksbank for their diligent and timely categorisation of all baseline radiographs using the Neer classification.

Finally, we thank the National Institute for Health Research Health Technology Assessment programme for its funding of the study and fortitude while patient recruitment was challenging and its congenial and effective staff who have often engaged with us in lengthy communication about various aspects of the study.

Contributions of authors

Dr Helen Handoll (Senior Research Fellow, Musculoskeletal Studies) contributed to the methods and reporting throughout the trial and wrote the first draft of the final report, incorporating specific reports from AK and BC, and all subsequent drafts.

Dr Stephen Brealey (Research Fellow, Trial Management) was the trial manager and as such was responsible for the management of the ProFHER trial throughout and for various aspects of trial design. He contributed to all drafts of the report.

Professor Amar Rangan (Clinical Professor, Consultant Orthopaedic Surgeon) was chief investigator and clinical lead and as such was responsible for the oversight of the conduct of the ProFHER trial and for various aspects of trial design. He contributed to all drafts of the report.

Ms Ada Keding (Research Fellow, Statistician) was the trial statistician during the last stage of the ProFHER trial and was responsible for finalising, executing and reporting the clinical effectiveness analysis. She commented on any statistical aspects of the report.

Mrs Belen Corbacho (Research Fellow, Health Economics) conducted and reported the economic evaluation. She commented on health economic aspects of the report.

Dr Laura Jefferson (Research Fellow, Trial Management) was responsible for co-ordination of the ProFHER trial and commented on all drafts of the report.

Dr Ling-Hsiang Chuang (Research Fellow, Health Economist) was responsible for writing the health economics analysis plan and commented on health economic aspects of the report.

Mrs Lorna Goodchild (Extended Scope Physiotherapist, Physiotherapy) was the physiotherapy lead for the ProFHER trial and was responsible for trial design and conduct relating to rehabilitation. She commented on all drafts of the report.

Dr Catherine Hewitt (Senior Research Fellow, Statistician) led the statistical component of the ProFHER trial, repeated the primary analysis and commented on all drafts of the report.

Professor David Torgerson (Director, YTU) was responsible for trial design and commented on all drafts of the report.

Publications

Handoll H, Brealey S, Rangan A, Torgerson D, Dennis L, Armstrong A, *et al.* Protocol for the ProFHER (PROximal Fracture of the Humerus: Evaluation by Randomisation) trial: a pragmatic multi-centre randomised controlled trial of surgical versus non-surgical treatment for proximal fracture of the humerus in adults. *BMC Musculoskelet Disord* 2009;**10**:140.

Handoll HH, Brealey S, Rangan A, on behalf of the ProFHER trial team. Proximal fractures of the humerus: the ProFHER trial. *Osteoporos Rev* 2010;**18**:19–21.

Handoll HHG, Goodchild L, Brealey S, Hanchard NCA, Jefferson L, Keding A, *et al.* Developing, delivering and documenting rehabilitation in a multi-centre randomised controlled surgical trial: experiences from the ProFHER trial. *Bone Joint Res* 2014;**3**:335–40.

Rangan A, Handoll H, Brealey S, Jefferson L, Keding A, Corbacho Martin B, *et al.* Surgical vs nonsurgical treatment of adults with displaced fractures of the proximal humerus: the PROFHER randomized clinical trial. *JAMA* 2015;**313**:1037–47.

References

1. Johnell O, Kanis JA. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Osteoporos Int* 2006;**17**:1726–33. <http://dx.doi.org/10.1007/s00198-006-0172-4>
2. Court-Brown CM, Caesar B. Epidemiology of adult fractures: a review. *Injury* 2006;**37**:691–7. <http://dx.doi.org/10.1016/j.injury.2006.04.130>
3. van Staa TP, Dennison EM, Leufkens HG, Cooper C. Epidemiology of fractures in England and Wales. *Bone* 2001;**29**:517–22. [http://dx.doi.org/10.1016/S8756-3282\(01\)00614-7](http://dx.doi.org/10.1016/S8756-3282(01)00614-7)
4. Office for National Statistics. *First Results on Population for England and Wales*. London: Office for National Statistics; 2002.
5. Lind T, Kroner K, Jensen J. The epidemiology of fractures of the proximal humerus. *Arch Orthop Trauma Surg* 1989;**108**:285–7. <http://dx.doi.org/10.1007/BF00932316>
6. Kim SH, Szabo RM, Marder RA. Epidemiology of humerus fractures in the United States: nationwide emergency department sample, 2008. *Arthritis Care Res* 2012;**64**:407–14. <http://dx.doi.org/10.1002/acr.21563>
7. Court-Brown CM, Garg A, McQueen MM. The epidemiology of proximal humeral fractures. *Acta Orthop Scand* 2001;**72**:365–71. <http://dx.doi.org/10.1080/000164701753542023>
8. Palvanen M, Kannus P, Niemi S, Parkkari J. Epidemiology of proximal humeral fractures. *Clin Orthop* 2006;**442**:87–92. <http://dx.doi.org/10.1097/01.blo.0000194672.79634.78>
9. Lubbeke A, Stern R, Grab B, Herrmann F, Michel J-P, Hoffmeyer P. Upper extremity fractures in the elderly: consequences on utilization of rehabilitation care. *Aging Clin Exp Res* 2005;**17**:276–80. <http://dx.doi.org/10.1007/BF03324610>
10. Health and Social Care Information Centre. *Hospital Episode Statistics*. URL: www.hscic.gov.uk/hes (accessed 29 May 2013).
11. Hodgson SA, Mawson SJ, Saxton JM, Stanley D. Rehabilitation of two-part fractures of the neck of the humerus (two-year follow-up). *J Shoulder Elbow Surg* 2007;**16**:143–5. <http://dx.doi.org/10.1016/j.jse.2006.06.003>
12. Olsson C, Petersson C, Nordquist A. Increased mortality after fracture of the surgical neck of the humerus. *Acta Orthop Scand* 2003;**74**:714–17. <http://dx.doi.org/10.1080/00016470310018252>
13. Melton LJ, Gabriel SE, Crowson CS, Torteson AN, Johnell O, Kanis JA. Cost-equivalence of different osteoporotic fractures. *Osteoporos Int* 2003;**14**:383–8. <http://dx.doi.org/10.1007/s00198-003-1385-4>
14. Ohsfeldt RL, Borisov NN, Sheer RL. Fragility fracture-related direct medical costs in the first year following a nonvertebral fracture in a managed care setting. *Osteoporos Int* 2006;**17**:252–8. <http://dx.doi.org/10.1007/s00198-005-1993-2>
15. Bell JE, Leung BC, Spratt KF, Koval KJ, Weinstein JD, Goodman DC, et al. Trends and variation in incidence, surgical treatment, and repeat surgery of proximal humeral fractures in the elderly. *J Bone Joint Surg Am* 2011;**93**:121–31. <http://dx.doi.org/10.2106/JBJS.I.01505>
16. Polinder S, Lordens GL, Panneman MJ, Eygendaal D, Patka P, Den Hartog D, et al. Trends in incidence and costs of injuries to the shoulder, arm and wrist in the Netherlands between 1986 and 2008. *BMC Public Health* 2013;**13**:531. <http://dx.doi.org/10.1186/1471-2458-13-531>

17. Court-Brown CM, Bugler KE, Clement ND, Duckworth AD, McQueen MM. The epidemiology of open fractures in adults. A 15 year review. *Injury* 2012;**43**:891–7. <http://dx.doi.org/10.1016/j.injury.2011.12.007>
18. Neer CS 2nd. Displaced proximal humeral fractures. Part I. Classification and evaluation. *J Bone Joint Surg Am* 1970;**52**:1077–89.
19. Sidor ML, Zuckerman JD, Lyon T, Koval K, Cuomo F, Schoenberg N. The Neer classification system for proximal humerus fractures. An assessment of interobserver reliability and intraobserver reproducibility. *J Bone Joint Surg Am* 1993;**75**:1745–50.
20. Codman EA. *The Shoulder: Rupture of the Supraspinatus Tendon and Other Lesions in or about the Subacromial Bursa*. Boston, MA: Thomas Todd; 1934.
21. Carofino BC, Leopold SS. Classifications in brief: the Neer classification for proximal humerus fractures. *Clin Orthop* 2013;**417**:39–43. <http://dx.doi.org/10.1007/s11999-012-2454-9>
22. Brorson S, Hrobjartsson A. Training improves agreement among doctors using the Neer system for proximal humeral fractures in a systematic review. *J Clin Epidemiol* 2008;**61**:7–16. <http://dx.doi.org/10.1016/j.jclinepi.2007.04.014>
23. Browner BD, Jupiter JB, Levine AM, Trafton AM, editors. *Skeletal Trauma: Basic Science, Management and Reconstruction. Volume 2*. Oxford: WB Saunders; 1991.
24. Bunker TD, Colton CL, Webb JK. *Frontiers in Fracture Management*. London: Martin Dunitz; 1989.
25. Hodgson S. Proximal humerus fracture rehabilitation. *Clin Orthop* 2006;**442**:131–8.
26. Handoll HH, Gibson JN, Madhok R. Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev* 2003;**4**:CD000434.
27. Kristiansen B, Kofoed H. Transcutaneous reduction and external fixation of displaced fractures of the proximal humerus. A controlled clinical trial. *J Bone Joint Surg Br* 1988;**70**:821–4.
28. Stableforth PG. Four-part fractures of the neck of the humerus. *J Bone Joint Surg Br* 1984;**66**:104–8.
29. Zyto K, Ahrengart L, Sperber A, Tornkvist H. Treatment of displaced proximal humeral fractures in elderly patients. *J Bone Joint Surg Br* 1997;**79**:412–17. <http://dx.doi.org/10.1302/0301-620X.79B3.7419>
30. Handoll HH, Ollivere BJ, Rollins KE. Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev* 2012;**12**:CD000434.
31. Fjalestad T, Hole MO, Jorgensen JJ, Stromsoe K, Kristiansen IS. Health and cost consequences of surgical versus conservative treatment for a comminuted proximal humeral fracture in elderly patients. *Injury* 2010;**41**:599–605. <http://dx.doi.org/10.1016/j.injury.2009.10.056>
32. Olerud P, Ahrengart L, Ponzer S, Saving J, Tidermark J. Internal fixation versus nonoperative treatment of displaced 3-part proximal humeral fractures in elderly patients: a randomized controlled trial. *J Shoulder Elbow Surg* 2011;**20**:747–55. <http://dx.doi.org/10.1016/j.jse.2010.12.018>
33. Olerud P, Ahrengart L, Ponzer S, Saving J, Tidermark J. Hemiarthroplasty versus nonoperative treatment of displaced 4-part proximal humeral fractures in elderly patients: a randomized controlled trial. *J Shoulder Elbow Surg* 2011;**20**:1025–33. <http://dx.doi.org/10.1016/j.jse.2011.04.016>

34. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). *Am J Ind Med* 1996;**29**:602–8. [http://dx.doi.org/10.1002/\(SICI\)1097-0274\(199606\)29:6<602::AID-AJIM4>3.0.CO;2-L](http://dx.doi.org/10.1002/(SICI)1097-0274(199606)29:6<602::AID-AJIM4>3.0.CO;2-L)
35. Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. *J Bone Joint Surg Br* 1996;**78**:593–600.
36. Huttunen TT, Launonen AP, Pihlajamäki H, Kannus P, Mattila VM. Trends in the surgical treatment of proximal humeral fractures – a nationwide 23-year study in Finland. *BMC Musculoskelet Disord* 2012;**13**:261. <http://dx.doi.org/10.1186/1471-2474-13-261>
37. Helmy N, Hinterman B. New trends in the treatment of proximal humerus fractures. *Clin Orthop* 2006;**442**:100–8. <http://dx.doi.org/10.1097/01.blo.0000194674.56764.c0>
38. Lill H, Katthagen C, Hertel A, Gille J, Voigt C. All-arthroscopic intramedullary nailing of 2- and 3-part proximal humeral fractures: a new arthroscopic technique and preliminary results. *Arch Orthop Trauma Surg* 2012;**132**:641–7. <http://dx.doi.org/10.1007/s00402-011-1430-2>
39. Panesar SS, Shaerf DA, Mann BS, Malik AK. Patient safety in orthopaedics: state of the art. *J Bone Joint Surg Br* 2012;**94**:1595–7. <http://dx.doi.org/10.1302/0301-620X.94B12.30217>
40. Handoll H, Brealey S, Rangan A, Torgerson D, Dennis L, Armstrong A, *et al.* Protocol for the ProFHER (PROximal Fracture of the Humerus: Evaluation by Randomisation) trial: a pragmatic multi-centre randomised controlled trial of surgical versus non-surgical treatment for proximal fracture of the humerus in adults. *BMC Musculoskelet Disord* 2009;**10**:140. <http://dx.doi.org/10.1186/1471-2474-10-140>
41. Dawson J, Rogers K, Fitzpatrick R, Carr A. The Oxford shoulder score revisited. *Arch Orthop Trauma Surg* 2009;**129**:119–23. <http://dx.doi.org/10.1007/s00402-007-0549-7>
42. Jain NP, Jowett AJ, Clarke NM. Learning curves in orthopaedic surgery: a case for super-specialisation? *Ann R Coll Surg Engl* 2007;**89**:143–6. <http://dx.doi.org/10.1308/003588407X155798>
43. Ramsay CR, Grant AM, Wallace SA, Garthwaite PH, Monk AF, Russell IT. Statistical assessment of the learning curves of health technologies. *Health Technol Assess* 2001;**5**(12). <http://dx.doi.org/10.3310/hta5120>
44. Griffin XL, Costa ML, Parsons N, Smith N. Electromagnetic field stimulation for treating delayed union or non-union of long bone fractures in adults. *Cochrane Database Syst Rev* 2011;**4**:CD008471.
45. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop* 1987;**214**:160–4.
46. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF36). I. Conceptual framework and item selection. *Med Care* 1992;**30**:473–83. <http://dx.doi.org/10.1097/00005650-199206000-00002>
47. Dawson J, Hill G, Fitzpatrick R, Carr A. The benefits of using patient-based methods of assessment. *J Bone Joint Surg Br* 2001;**83**:877–82. <http://dx.doi.org/10.1302/0301-620X.83B6.11316>
48. Ware JE, Kosinski N, Keller SD. The 12-Item Short-Form Health Survey (SF-12): construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;**34**:220–33. <http://dx.doi.org/10.1097/00005650-199603000-00003>
49. EuroQol Group. EuroQol: a new facility for the measurement of health related quality of life. *Health Policy* 1990;**16**:199–208. [http://dx.doi.org/10.1016/0168-8510\(90\)90421-9](http://dx.doi.org/10.1016/0168-8510(90)90421-9)

50. Baker P, Nanda R, Goodchild L, Finn P, Rangan A. A comparison of the Constant and Oxford shoulder scores in patients with conservatively treated proximal humeral fractures. *J Shoulder Elbow Surg* 2008;**17**:37–41. <http://dx.doi.org/10.1016/j.jse.2007.04.019>
51. Lee K, Thompson S. Clustering by health professional in individually randomised trials. *BMJ* 2005;**330**:142–4. <http://dx.doi.org/10.1136/bmj.330.7483.142>
52. Department of Health. *Research Governance Framework for Health and Social Care*. 2nd edn. London: Department of Health; 2005. URL: www.gov.uk/government/publications/research-governance-framework-for-health-and-social-care-second-edition (accessed 16 December 2014).
53. British Orthopaedic Association. *The Care of Patients with Fragility Fracture ('Blue Book')*. London: British Orthopaedic Association; 2007.
54. Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ* 2010;**340**:c117. <http://dx.doi.org/10.1136/bmj.c117>
55. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. London: NICE; 2013.
56. Burton A, Billingham L, Bryan S. Cost-effectiveness in clinical trials: using multiple imputation to deal with incomplete cost data. *Clin Trials* 2007;**4**:154–61. <http://dx.doi.org/10.1177/1740774507076914>
57. Manca A, Palmer S. Handling missing data in patient level cost-effectiveness analysis alongside randomised clinical trials. *Appl Health Econ Health Policy* 2005;**4**:65–75. <http://dx.doi.org/10.2165/00148365-200504020-00001>
58. Lin DY, Feuer EJ, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data. *Biometrics* 1997;**53**:419–34. <http://dx.doi.org/10.2307/2533947>
59. Neer CS 2nd. Four-segment classification of proximal humeral fractures: purpose and reliable use. *J Shoulder Elbow Surg* 2002;**11**:389–400. <http://dx.doi.org/10.1067/mse.2002.124346>
60. White E. Developing the Allied Health Professions Research Network: building research capacity and capability. *Int J Ther Rehabil* 2012;**19**:601–2. <http://dx.doi.org/10.12968/ijtr.2012.19.11.601>
61. Carr A, Cooper C. The UKUFF trial and the NIHR comprehensive local research networks. *Shoulder Elbow* 2009;**1**:63–4. <http://dx.doi.org/10.1111/j.1758-5740.2009.00030.x>
62. Edwards P, Roberts I, Clarke M, DiGuseppi C, Pratap S, Wentz R, et al. Increasing response rates to postal questionnaires: systematic review. *BMJ* 2002;**324**:1–9. <http://dx.doi.org/10.1136/bmj.324.7347.1183>
63. Kristiansen B, Andersen UL, Olsen CA, Varmarken JE. The Neer classification of fractures of the proximal humerus. An assessment of interobserver variation. *Skeletal Radiol* 1988;**17**:420–2. <http://dx.doi.org/10.1007/BF00361661>
64. Siebenrock KA, Gerber C. The reproducibility of classification of fractures of the proximal end of the humerus. *J Bone Joint Surg Am* 1993;**75**:1751–5.
65. Brorson S, Bagger J, Sylvest A, Hobjartsson A. Improved interobserver variation after training of doctors in the Neer system. A randomised trial. *J Bone Joint Surg Br* 2002;**84**:950–4. <http://dx.doi.org/10.1302/0301-620X.84B7.13010>
66. Neer CS 2nd. Displaced proximal humeral fractures. Part II. Treatment of three-part and four-part displacement. *J Bone Joint Surg Am* 1970;**52**:1090–103.
67. Health and Social Care Information Centre. *Would PACS Have Happened Anyway?* URL: <http://systems.hscic.gov.uk/pacs/learn/different/myth> (accessed 31 October 2013).

68. Department of Health Informatics Directorate. *NHS Connecting for Health* (ceased 31 March 2013). URL: <http://webarchive.nationalarchives.gov.uk/20130502102046/http://www.connectingforhealth.nhs.uk/> (accessed 15 December 2014).
69. Ware JE Jr, Kosinski M, Turner-Bowker DM, Gandek B. *User's Manual for the SF-12v2® Health Survey (With a Supplement Documenting SF-12® Health Survey)*. Lincoln, RI: QualityMetric; 2009.
70. Drummond MF, O'Brien B, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press; 2005.
71. Torgerson DJ, Torgerson C. *Designing Randomised Trials in Health, Education and the Social Sciences: An Introduction*. Basingstoke: Palgrave Macmillan; 2008. <http://dx.doi.org/10.1057/9780230583993>
72. Brooks R. EuroQol: the current state of play. *Health Policy* 1996;**37**:53–72. [http://dx.doi.org/10.1016/0168-8510\(96\)00822-6](http://dx.doi.org/10.1016/0168-8510(96)00822-6)
73. Dolan P, Gudex C, Kind P, Williams A. *A Social Tariff for EuroQol: Results from a UK General Population Survey*. York: University of York, Centre for Health Economics; 1995.
74. Olerud P, Tidermark J, Ponzer S, Ahrengart L, Bergström G. Responsiveness of the EQ-5D in patients with proximal humeral fractures. *J Shoulder Elbow Surg* 2011;**20**:1200–6. <http://dx.doi.org/10.1016/j.jse.2011.06.010>
75. Billingham L, Abrams KR, Jones DR. Methods for the analysis of quality-of-life and survival data in health technology assessment. *Health Technol Assess* 1998;**3**(10).
76. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 2005;**14**:487–96. <http://dx.doi.org/10.1002/hec.944>
77. Curtis L. *Unit Costs of Health and Social Care 2012*. Canterbury: PSSRU, University of Kent; 2012.
78. Department of Health. *NHS Reference Costs 2011/12*. 2012. URL: www.dh.gov.uk/health/2012/11/2011-12-reference-costs/ (accessed 10 October 2013).
79. Joint Formulary Committee. *British National Formulary*. 66 ed. London: BMJ Group and Pharmaceutical Press; 2013.
80. Little RJ, Rubin DB. *Statistical Analysis with Missing Data*. New York, NY: Wiley; 1987.
81. White IR, Horton NJ, Carpenter J. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ* 2011;**342**:d40. <http://dx.doi.org/10.1136/bmj.d40>
82. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;**30**:377–99. <http://dx.doi.org/10.1002/sim.4067>
83. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York, NY: Wiley; 1987. <http://dx.doi.org/10.1002/9780470316696>
84. Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health Econ* 2004;**13**:461–75. <http://dx.doi.org/10.1002/hec.843>
85. Manca A, Lambert PC, Sculpher M, Rice N. Cost-effectiveness analysis using data from multinational trials: the use of bivariate hierarchical modelling. *Med Decis Making* 2007;**27**:471–90. <http://dx.doi.org/10.1177/0272989X07302132>
86. Johannesson M, Weinstein MC. On the decision rules of cost-effectiveness analysis. *J Health Econ* 1993;**12**:459–67. [http://dx.doi.org/10.1016/0167-6296\(93\)90005-Y](http://dx.doi.org/10.1016/0167-6296(93)90005-Y)

87. Stinnett AA, Mullahy J. Net health benefits a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Med Decis Making* 1998;**18**:S68–80. <http://dx.doi.org/10.1177/0272989X9801800209>
88. Claxton K. Exploring uncertainty in cost-effectiveness analysis. *Pharmacoeconomics* 2008;**26**:781–98. <http://dx.doi.org/10.2165/00019053-200826090-00008>
89. Efron B, Tibshirani R. *An Introduction to the Bootstrap*. New York, NY: Chapman & Hall; 1993. <http://dx.doi.org/10.1007/978-1-4899-4541-9>
90. Briggs A. Economics notes: handling uncertainty in economic evaluation. *BMJ* 1999;**319**:120. <http://dx.doi.org/10.1136/bmj.319.7202.120>
91. Van de Water AT, Shields N, Davidson M, Evans M, Taylor NF. Reliability and validity of shoulder function outcome measures in people with a proximal humeral fracture. *Disabil Rehabil* 2014;**36**:1072–9. <http://dx.doi.org/10.3109/09638288.2013.829529>
92. Hanson B, Neidenbach P, de Boer P, Stengel D. Functional outcomes after nonoperative management of fractures of the proximal humerus *J Shoulder Elbow Surg* 2009;**18**:612–21. <http://dx.doi.org/10.1016/j.jse.2009.03.024>
93. Murray IR, Amin AK, White TO, Robinson CM. Proximal humeral fractures: current concepts in classification, treatment and outcomes. *J Bone Joint Surg Br* 2011;**93**:1–11. <http://dx.doi.org/10.1302/0301-620X.93B1.25702>
94. Brorson S, Olsen BS, Frich LH, Jensen SL, Sorensen AB, Krogsgaard MR, et al. Surgeons agree more on treatment recommendations than on classification of proximal humeral fractures. *BMC Musculoskelet Disord* 2012;**13**:114. <http://dx.doi.org/10.1186/1471-2474-13-114>
95. Boons HW, Goosen JH, Grinsven S, Susante JL, Loon CJ. Hemiarthroplasty for humeral four-part fractures for patients 65 years and older: a randomized controlled trial. *Clin Orthop* 2012;**470**:3483–91. <http://dx.doi.org/10.1007/s11999-012-2531-0>
96. Launonen AP, Lepola V, Flinkkila T, Strandberg N, Ojanpera J, Rissanen P, et al. Conservative treatment, plate fixation, or prosthesis for proximal humeral fracture. A prospective randomized study. *BMC Musculoskelet Disord* 2012;**13**:167. <http://dx.doi.org/10.1186/1471-2474-13-167>
97. Thanasis C, Kontakis G, Angoules A, Limb D, Giannoudis P. Treatment of proximal humerus fractures with locking plates: a systematic review. *J Shoulder Elbow Surg* 2009;**18**:837–44. <http://dx.doi.org/10.1016/j.jse.2009.06.004>
98. Brorson S, Olsen BS, Frich LH, Jensen SL, Johannsen HV, Sorensen AK, et al. Effect of osteosynthesis, primary hemiarthroplasty, and non-surgical management for displaced four-part fractures of the proximal humerus in elderly: a multi-centre, randomised clinical trial. *Trials* 2009;**10**:51. <http://dx.doi.org/10.1186/1745-6215-10-51>
99. ClinicalTrials.gov. *Operative Versus Non Operative Treatment of Proximal Humerus (Shoulder Joint) Fractures*. URL: <http://clinicaltrials.gov/show/NCT00818987> (accessed 2 December 2012).
100. ClinicalTrials.gov. *Treatment of Comminuted Fractures of the Proximal Humerus. A Randomised, Controlled Study*. <http://clinicaltrials.gov/ct2/show/record/NCT00999193> (accessed 2 December 2012).
101. Den Hartog D, van Lieshout EM, Tuinebreijer WE, Polinder S, Van Beeck EF, Breederveld RS, et al. Primary hemiarthroplasty versus conservative treatment for comminuted fractures of the proximal humerus in the elderly (ProCon): a multicenter randomized controlled trial. *BMC Musculoskelet Disord* 2010;**11**:97. <http://dx.doi.org/10.1186/1471-2474-11-97>

102. Acevedo DC, Mann T, Abboud JA, Getz C, Baumhauer JF, Voloshin I. Reverse total shoulder arthroplasty for the treatment of proximal humeral fractures: patterns of use among newly trained orthopedic surgeons. *J Shoulder Elbow Surg* 2014;**23**:1363–7. <http://dx.doi.org/10.1016/j.jse.2014.01.005>
103. Johansen A, Wakeman R, Boulton C, Plant F, Roberts J, Williams A. *The National Hip Fracture Database. National Report 2013*. URL: [www.nhfd.co.uk/20/hipfractureR.nsf/luMenuDefinitions/CA920122A244F2ED802579C900553993/\\$file/NHFD%20Report%202013.pdf?OpenElement](http://www.nhfd.co.uk/20/hipfractureR.nsf/luMenuDefinitions/CA920122A244F2ED802579C900553993/$file/NHFD%20Report%202013.pdf?OpenElement) (accessed 13 March 2015).
104. Brorson S, Eckardt H, Audigé L, Rolauffs B, Bahrs C. Translation between the Neer- and the AO/OTA-classification for proximal humeral fractures: do we need to be bilingual to interpret the scientific literature? *BMC Res Notes* 2013;**6**:69. <http://dx.doi.org/10.1186/1756-0500-6-69>
105. Castillo RC, Bosse MJ, MacKenzie EJ, Patterson BM, the LEAP study group. Impact of smoking on fracture healing and risk of complications in limb-threatening open tibia fractures. *J Orthop Trauma* 2005;**19**:151–7. <http://dx.doi.org/10.1097/00005131-200503000-00001>
106. Schmitz MA, Finnegan M, Natarajan R, Rajeshwari MA, Champine J. Effect of smoking on tibial shaft fracture healing. *Clin Orthop* 1999;**365**:184–200. <http://dx.doi.org/10.1097/00003086-199908000-00024>
107. Møller AM, Pedersen T, Villebro N, Munksgaard A. Effect of smoking on early complications after elective orthopaedic surgery. *J Bone Joint Surg Br* 2003;**85-B**:178–81. <http://dx.doi.org/10.1302/0301-620X.85B2.13717>
108. Lin DY. Regression analysis of incomplete medical cost data. *Stat Med* 2003;**22**:1181–200. <http://dx.doi.org/10.1002/sim.1377>
109. Royston P. Multiple imputation of missing values: update of ice. *Stata J* 2005;**5**:527–36.

















Appendix 1 Hospital Episode Statistics for England: inpatient statistics (2003/4 to 2011/12) – totals and proximal humeral fracture

Group	FCEs, <i>n</i>	Admissions, <i>n</i>	Male, <i>n</i>	Emergency, <i>n</i>	Waiting list, <i>n</i>	Time waited (days)		Length of stay (days)		Mean age (years)	Age (years), <i>n</i>				Day case, <i>n</i>	FCE bed-days, <i>n</i>
						Mean	Median	Mean	Median		0–14	15–59	60–74	75+		
2011–12																
Total	17,465,425	15,019,396	7,754,620	5,242,839	5,541,731	52	35	5.3	1	51	1,990,144	7,438,486	3,864,690	4,098,361	5,924,107	48,631,585
S42.2	15,918	12,077	4658	9708	1484	9	5	8.2	2	71	399	3135	3873	8495	107	97,489
2010–11																
Total	17,269,882	14,890,844	7,628,685	5,287,032	5,431,710	50	35	5.5	1	51	1,992,277	7,463,084	3,773,052	3,982,390	5,691,706	51,210,196
S42.2	15,276	11,761	4640	9590	1299	9	5	8.3	2	71	375	3058	3743	8083	53	97,774
2009–10																
Total	16,806,196	14,537,712	7,408,085	5,177,887	5,344,982	50	34	5.6	1	51	1,939,193	7,333,114	3,642,944	3,837,988	5,474,889	51,493,494
S42.2	15,359	11,824	4770	9777	1238	9	5	8.8	3	70	360	3132	3694	8160	55	104,935
2008–9																
Total	16,232,579	14,152,692	7,145,742	5,010,670	5,167,989	48	35	5.7	1	50	1,901,736	7,132,442	3,500,406	3,637,446	5,220,516	51,841,443
S42.2	13,396	10,474	4163	8696	1138	9	5	8.7	3	69	394	3003	3180	6806	62	93,217
2007–8																
Total	15,359,062	13,479,828	6,721,648	4,753,368	4,862,278	60	42	5.7	1	50	1,840,024	6,929,479	3,241,986	3,314,849	4,766,383	48,267,447
S42.2	12,205	9648	3637	8100	980	10	5	9.2	3	69	419	2669	2792	6306	54	88,403
2006–7																
Total	14,784,581	12,976,273	6,483,429	4,700,017	4,550,689	73	49	6.3	2	49	1,791,408	6,724,391	3,066,605	3,174,676	4,373,390	50,080,570
S42.2	11,328	8946	3568	7510	859	13	5	10.0	3	69	406	2536	2455	5918	24	89,683
2005–6																
Total	14,423,506	12,678,628	6,303,012	4,659,054	4,368,056	78	51	6.6	2	49	1,764,562	6,585,487	2,949,144	3,095,729	4,113,304	52,920,218
S42.2	10,581	8408	3235	7152	666	13	5	10.7	4	69	389	2418	2308	5464	18	90,112
2004–5																
Total	13,706,765	12,102,006	5,983,455	4,428,680	4,187,619	84	52	7.1	2	49	1,719,476	6,220,473	2,801,430	2,937,276	3,847,632	54,554,697
S42.2	9378	7568	2926	6416	630	12	5	11.7	4	68	385	2134	2063	4790	26	87,806
2003–4																
Total	13,174,480	11,699,163	5,759,916	4,158,734	4,227,180	95	50	7.4	2	49	1,674,944	5,980,530	2,698,407	2,789,703	3,757,476	54,619,610
S42.2	7848	6427	2498	5479	540	12	5	12.1	5	68	396	1757	1643	4050	22	77,828

S42.2 – Fracture of Upper End of Humerus. Source: Health & Social Care Information Centre. Hospital Episode Statistics. URL: www.hscic.gov.uk/hes (accessed 15 February 2013).

S42.2 – Fracture of Upper End of Humerus. Source: Health & Social Care Information Centre. Hospital Episode Statistics. URL: www.hscic.gov.uk/hes (accessed 15 February 2013).

Appendix 2 Neer classification for proximal humeral fractures

	2 PART	3 PART	4 PART	
Anatomical Neck	 2			Minimal Displacement  1
Surgical Neck	 3			
Greater Tuberosity	 4	 8	 12	Articular Surface  15  16
Lesser Tuberosity	 5	 9		
Fracture-Dislocation Anterior	 6	 10	 13	
Posterior	 7	 11	 14	

This figure shows the 16 categories of the Neer classification. A fracture is considered displaced if one or more of the four segments are displaced > 1 cm or angulated $> 45^\circ$. The figure is modified from Neer¹⁸ with permission from the *Journal of Bone and Joint Surgery* and from Brorson *et al.*¹⁰⁴ with permission.

Appendix 3 The ProFHER trial protocol version 7.0 (19 May 2009)

Please note: the protocol published in this appendix is the final version of the protocol and is reproduced again as is. Therefore, the appendices listed are not included and those cited do not correspond with the appendices in the main report.

PROximal Fractures of the Humerus: Evaluation by Randomisation – The PROFHER Trial

This protocol describes a UK multi-centre randomised controlled trial of surgical versus non-surgical treatment for the majority of displaced fractures of the proximal humerus in adults.

The NHS R&D Health Technology Assessment Programme (HTA) is the sole external funder of this trial. This protocol is derived from the detailed project description of the HTA funding application entitled 'Pragmatic multi-centre randomised trial of surgical versus non-surgical treatment for proximal fracture of the humerus in adults' [HTA: 06/404/53].

The PROFHER trial is sponsored by the University of Teesside. Trial management is primarily by the York Trials Unit (YTU), University of York. This trial has received endorsement by the British Elbow and Shoulder Society.

Summary of planned investigation

Research objectives

Our primary aim is to obtain reliable evidence of effectiveness and cost-effectiveness of basic treatment strategies for the majority of displaced fractures of the proximal humerus in adults. Hence, we plan to undertake a pragmatic randomised clinical trial (RCT) of surgical versus non-surgical treatment of displaced proximal humeral fractures involving the surgical neck in adults.

Study population

After a period of preparation and upon MREC (Multicentre Research Ethics Committee) approval, recruitment into the RCT will be over an 18 month period from at least 18 NHS trauma centres; most of which are either already committed or have expressed keen interest in contributing to the trial, and thus identify with the study aims. We will promote a minimum recruitment rate of one patient per centre per month. Allowing for conservative estimate of recruitment success and for drop-outs we propose to recruit and randomise 250 patients over an 18 month period. We anticipate data from 200 patients followed up for 2 years. The HTA have accepted our proposal for an assessment point after 10 months of recruitment to determine if recruitment into the RCT is sufficient (set at 88 patients) to justify its continuation after 18 months.

Integral to the trial, and to inform on trial recruitment and applicability of trial results as well as to conform to trial-reporting standards, we plan a systematic prospective collection of key patient data for those meeting the main inclusion criteria (age 16 or over, presenting within 3 weeks of injury, with displaced fractures of the proximal humerus involving the surgical neck) for the RCT but who were not included. These will consist of three categories of patients: those that were excluded because they met one or more of the listed trial exclusion criteria, those who did not consent and those where there was a 'protocol violation' (reflecting lack of surgeon equipoise). The information collected for these patients will include data on patient preference, surgeon's advised treatment and the agreed treatment. In particular, we will

stress the importance of the systematic prospective collection of key patient data and reasons for exclusion of otherwise eligible patients. At the end of the 10 month 'feasibility' study, we anticipate 1400 potentially eligible or eligible patients would not be included in the randomised trial; should recruitment continue, we anticipate 2150 such patients.

Background

Proximal humeral fractures account for approximately 4–5% of all fractures. Their incidence rapidly increases with age, and women are affected over twice as often as men. Similar to other primarily osteoporotic fractures, the incidence of these fractures is increasing. Palvanen *et al.* 2006 found a three fold increase over a 33 year period in the incidence of proximal humeral fractures resulting from low-energy trauma in people aged 60 and above.

A large prospective epidemiology study (Court-Brown *et al.* 2001) found that around half of these fractures (51%) are displaced, when assessed according to the criteria of Neer's classification system (Neer 1970): one or more parts of the fractured bone are displaced by more than one centimetre, or angulated more than 45 degrees. Court-Brown *et al.* 2001 found that the largest groups of displaced fractures were 2 part surgical neck fractures (28% of the whole population), followed by 3 part greater tuberosity and surgical neck fractures (9%). Four part fractures without fracture dislocation were around 2% of the total. These figures are consistent with estimates from several members of the trial group.

Recent systematic reviews (Handoll *et al.* 2003; Misra *et al.* 2001), one of which was updated in 2007 (Handoll *et al.* 2003), have found a lack of evidence from randomised controlled trials (RCTs) to inform management decisions for proximal humeral fractures. In particular, there were only three completed RCTs comparing surgery with conservative treatment. All were small studies (numbers randomised: 30, 32, 40) with flawed methodology. Both reviews (Handoll *et al.* 2003; Misra *et al.* 2001) concluded that it was unclear whether or not operative intervention, even for specific fracture types, would produce consistently better long-term outcomes.

It is also clear from the literature, confirmed by an informal survey of the treatment provided by several UK centres, that there is great variation in the treatment of these fractures, both in basic (the use of surgery) and specific (type of implants and surgical technique; non-surgical management (Hodgson 2006) and rehabilitation packages) terms. Additionally, technology is changing all the time with various pressures towards early implementation.

The above findings point to a clear need to get reliable evidence to inform practice, and crucially to establish whether or not there is a role for operative intervention for the common types of acute displaced fractures of the proximal humerus. This is the focus of this trial.

Research methods

As indicated, we intend to undertake a pragmatic randomised clinical trial evaluating the effectiveness and cost-effectiveness of surgical intervention versus standard conservative therapy for the treatment of the majority of displaced (all involving the surgical neck) proximal humeral fractures in adults. This RCT includes the systematic collection of reasons for non-inclusion of eligible patients who were not recruited into the trial, and their baseline characteristics, treatment preferences and intended treatment.

Underpinning our approach are two key issues:

- There is a general dearth of reliable evidence to inform on the use of surgery (definitive treatment) for patients with these fractures.

- There are known difficulties in recruitment and particularly patient (and surgeon) preferences. Based on experience from previous studies, some abandoned, in this field it is anticipated that a large proportion of eligible individuals are likely to refuse to be randomised because they (or their surgeons) will have a strong preference for one of the study interventions; generally conservative in the case of patients. On discussions with orthopaedic surgeons, lack of clinical equipoise, which is another important barrier to performing surgical RCTs (Solomon *et al.* 1995), is anticipated to be less of an issue here. Because of these strong preferences it is likely that patients recruited into any RCT will be a highly selected group, which may threaten the external validity of our study. Thus, collecting key data for all patients eligible for the RCT will allow us to set our randomised results within the context of the whole patient population and give some pointers to the applicability of the results of the study.

How are the results of the trial to be used/interpreted?

The trial aims to establish whether or not surgery yields superior results to non-surgical treatment. As detailed below, our protocol emphasises standardised protocols and care pathways throughout, comparable and sufficient expertise of care providers and that the surgeon uses established techniques with which they are already familiar. Any questions over whether or not the use of other surgical methods, perhaps new methods, would give different results are countered by two considerations. Firstly, there is an absence of robust evidence to inform best surgical methods. Secondly, and arguably, the avoidance of 'learning curves' and the reliance on surgeon's competence is more representative of best surgical treatment.

Brief details of the proposed practical arrangements for trial recruitment and allocating participants to trial groups

A detailed generic scheme of the recruitment process will be devised for adoption according to local circumstances in the participating centres. At radiological review by the surgeon or their nominated deputy, a trial eligibility form (see *Appendix 1*) will be completed for any patient who meets the trial inclusion criteria. For ineligible patients, the surgeon is asked to indicate what treatment they would advise for the patient before the form is sent to the York Trials Unit. Those patients who the surgeon indicates as eligible for the trial will be invited to take part in the trial and the site-specific patient consent process is initiated. For non-consenting patients, this fact will be indicated on the Consent status form (see *Appendix 2*), where the surgeon or their deputy is asked to indicate their advised treatment, the patient's preferred treatment (if any) is completed.

Once patients have given consent (see *Appendix 3*), the recruiting clinician will complete the baseline data form (see *Appendix 4*) and then contact the York Trials Unit, either by telephone or via the internet, to access a secure randomisation service. This will ensure immediate and unbiased allocation of treatment.

Proposed methods for avoidance of bias and to ensure validity

Randomisation eliminates selection bias: there are, however, other forms of bias we will guard against. We will also take measures to ensure the external validity of trial results. We will undertake the following:

- Adherence to local guidelines for radiographic assessment will be actively promoted. If not stipulated already, we will encourage the use of the full shoulder trauma series (Neer 2002). Documentation including a power point presentation illustrating the full trauma series will be made available as part of the trial materials. A minimum of two x-ray views/projections is required for the assessment of study eligibility.
- At the end of the recruitment period, there will be scrutiny and categorisation based on the Neer classification system, using pre-prepared forms, of the baseline X-rays of all randomised patients. This will be done by an independent panel of musculoskeletal radiologists or orthopaedic surgeons who have experience with the Neer classification (Neer 1970). Copies of X-rays will be prepared beforehand to ensure they are anonymised. On an on-going basis during trial recruitment there will also be a review of the quality of the copies of X-ray images for each trial participant provided by trial centres.

This is to ensure that at the end of trial recruitment the images are of sufficient quality for the independent panel to assess and classify the fractures. For the X-rays of the first five participants at each centre, independent assessment of the quality of images will be done by three orthopaedic surgeons, one of whom will be the Chief Investigator. Assessment of the X-rays of subsequent participants of each centre will be done by one orthopaedic surgeon (the Chief Investigator), who if he has concerns will ask the two other independent surgeons for their comments.

- Clear entry criteria, including checks at randomisation, will reduce inappropriate entry into the RCT.
- We will endeavour to provide a consistent approach to recruitment and obtaining informed consent by providing an unbiased account of the study to eligible participants using a specially produced information sheet. These materials will be produced in collaboration with service users.
- Concealment of treatment allocation prior to trial entry will be ensured by use of an independent telephone randomisation service, as provided by the University of York's Trials Unit. After an initial, prespecified, period of randomisation, stratified by the presence or not of a tuberosity fracture and with a prespecified block size, randomisation will then be performed using a computer generated minimisation programme. The minimisation factors will be fractures involving either tuberosity and centre.
- We will emphasise good practice and standardised protocols and care pathways throughout, and comparable and sufficient expertise of care providers. Surgery for these types of fractures are usually carried out by consultants; this has been confirmed by an informal survey of the centres initially included in our study. We will attempt to minimise 'learning curve' issues for the surgical interventions by allowing the surgeon to use techniques with which they are familiar, but prohibiting the introduction of radically new or experimental methods during the recruitment period.
- We will encourage the prescription of comparable care, including rehabilitation programmes, such that any substantive departures from the norm would reflect the special requirements of a specific intervention. Consensus guidelines for non-surgical treatment and rehabilitation for both groups are being prepared by rehabilitation specialists and will be circulated for comment and input. These will form part of the trial materials and we will request details of where the prescribed treatment differs substantively from the standardised protocols. (See Notes added in clarification below.)

Notes added for clarification

1. Upon discussion, the proposed consensus guidelines for non-surgical treatment were considered inappropriate in the context of a pragmatic trial and the lack of evidence to inform practice. It was decided that the onus should be on the provision of good standard care and that our approach would be to indicate both verbally and in the site manual that we would anticipate initial care to comprise sling immobilisation for about 3 weeks or for as long as the treating clinician deemed necessary and active early rehabilitation. We considered that written information to advise patients during sling immobilisation was needed and should be provided to all eligible patients for the trial. We provide a generic document to be adopted by the hospital should a suitable document not already be available locally.
2. We stipulate that physiotherapy should be provided equally to both treatment groups. A consensus protocol giving basic treatment guidelines has been devised. Although, deviation from the protocol is allowed and expected, we stipulate that electrotherapy (except TENS) is not used, and point to an absence of evidence for these modalities as well as endorsement via a consultation process with specialist shoulder physiotherapists. We will promote the need to encourage home exercises, but decided not to provide generic information leaflets illustrating exercises for home use by patients. Instead we will check that physiotherapists either provide these already or access a standard web-based facilities to generate 'bespoke' exercise sheets.
3. We are prospectively collecting details of rehabilitation treatment which will also allow the detection of substantial differences from the physiotherapy protocol.

- We shall follow the CONSORT guidelines for considering and reporting RCTs (Moher *et al.* 2001).
- For instance, if eligible patients decline to be randomised, then this is refusal to consent (as per CONSORT); if surgeons choose not to randomise an eligible patient then this is a break in protocol (protocol violation) which must be recorded along with a reason.
- Intention to treat analyses will be undertaken as the primary analysis in the RCT.
- Active and systematic follow-up of all randomised participants at 3, 6, 12 and 24 months is planned.
- This will include pre-notification letters as well as the use of reminders after 2 and 4 weeks. For 6, 12 and 24 month follow-ups, there will be an option for completion of an abridged questionnaire via telephone after 6 weeks. We will also include an unconditional incentive payment of £5 to maximise the 12 and 24 month follow-up.
- As far as possible, all participants will be followed up for any unplanned events. Their hospital notes will include a reminder to notify of relevant subsequent treatment/events and they will be flagged for mortality. With the participant's permission, letters will be sent to their General Practitioners to inform of participation (see Appendix 6). *Participant's permission will also be sought to allow us to ask their GP to provide the participant's contact details should there be problems contacting them directly.*
- There will be independent data entry, processing and analysis. Aside from accrual and whole populations baseline statistics, interim results will not be made available to the trial investigators or associates in the participant centres.

Data collection on all potentially eligible patients

In addition to the systematic collection of basic baseline data for those eligible for the RCTs but who did not consent or where there was a protocol violation (reflecting lack of surgeon equipoise) to satisfy the requirements of CONSORT, we will collect data on patient-preferred and intended management. To complete the CONSORT flow diagram, we will collect the baseline data and reasons for ineligibility for ineligible adults presenting in the recruiting centres with the study fractures: see inclusion criteria.

Pilot study

The study will initially be set up in Teesside for training, and piloting materials and procedures.

Ethical arrangements

MREC (Multicentre Research Ethics Committee) approval has been obtained from York Research Ethics Committee (reference number 08/H1311/12). Consequently, separate submissions of the protocol are being submitted to the ethical committee of each centre. The approval of the HTA funding and MREC is being raised in support of applications to preclude unsatisfactory local variation.

Risks and anticipated benefits for trial participants and society, including how benefits justify risks

In the context of the lack of robust evidence to determine the best treatment for patients with these fractures, the risks are not increased through trial and/or study participation. Measures, such as our emphasis on good practice and standardised protocols/care pathways throughout, taken by us are indeed likely to reduce risk and could bring additional benefits. We will emphasise the importance of surgeons performing operations with which they are familiar and undertake on a regular basis. We will also stress the importance of competence in conservative methods, principally rehabilitation. We will stipulate that radically different novel devices or methods are not introduced at each centre for the duration of the trial; any necessary training will focus on already established methods. We will adhere to the good clinical research practice guidelines (MRC and Research Governance Framework). Though we will perform active and systematic follow-up, the timing of early follow-up parallels usual clinical practice and we will not place an undue burden on participants. Our adoption of the Oxford Shoulder Score, Euroqol (EQ-5D) and SF12, all of which are self completion questionnaires, avoids the need for participants to specially return for clinical follow-up assessments at 6, 12 and 24 months.

Ultimately, the RCT will provide evidence that should make a major contribution to the future management of these fractures. It is conceivable that the heightened awareness of and focus on these fractures in each centre could directly benefit study and trial participants: for instance, there is some evidence of improved outcomes of participants participating in RCTs.

Informing potential trial participants of possible benefits and known risks

A participant information sheet for the study has been compiled, with involvement of service users (see *Appendix 5*). This aims to give a balanced account of the possible benefits and known risks of the interventions under test. It states explicitly that quality of care will not be compromised if the participant decides to a) not enter the trial or b) withdraw their consent. We believe that the information provided is clear and easily understandable to the reader. Surgery is often performed the following day and therefore the information leaflet will be made available for the potential participant to allow them time to make an informed decision. Where possible, translations of the participant information sheet will be provided for non-English speakers.

Obtaining informed consent from participants whenever possible or proposed action where fully informed consent is not possible

Written informed consent will be obtained from all participants. Where possible, provision of interpreters will be made on a local basis for non-English speaking participants. Where local interpreters are unavailable, we will consider using the National Interpreting Service. No participant will be entered into the RCT without informed consent.

Proposed time period for retention of relevant trial documentation

Minimum 20 years.

Participants: planned inclusion/exclusion criteria

Inclusion criteria

Adults (aged 16 or above) presenting to the participating trauma centre within 3 weeks of their injury with a radiologically confirmed displaced fracture of the proximal humerus involving the surgical neck. This should include all 2 part surgical neck fractures; 3 part (including surgical neck) and 4 part fractures of proximal humerus (Neer Classification). It may also include displaced surgical neck fractures that do not meet the exact displacement criteria of the Neer classification (1 cm or/and 45° angulation of displaced parts) where this reflects an individual surgeon's equipoise (e.g., whether or not the surgical neck fracture should be treated surgically).

Exclusion criteria

- Associated dislocation of the injured shoulder joint
- Open fracture
- Mentally incompetent patient: unable to understand trial procedure or instructions for rehabilitation; significant mental impairment that would preclude compliance with rehabilitation and treatment advice
- Co-morbidities precluding surgery/anaesthesia
- A clear indication for surgery such as severe soft-tissue compromise requiring surgery/emergency treatment (nerve injury/dysfunction)
- Multiple injuries: same limb fractures; other upper limb fractures
- Pathological fractures (other than osteoporotic) & terminal illness
- Participant not resident in trauma-centre catchment area.

Sample population

This will be all adults (aged 16 or above) presenting within 3 weeks of injury with fracture types listed in the inclusion criteria. Ineligible patients will be defined as those who are excluded for reasons given in the exclusion criteria. All those who meet the above criteria will be termed eligible patients. Some patients still may not be included in the RCT, for instance due to lack of patient consent (patient has strong preference for specific treatment option or refuses randomisation) or because the surgeon considers one of the treatment options is strongly indicated for reasons other than above.

Interventions

Each centre participating in this trial has to agree to forgo the introduction of radically novel and experimental interventions for these fractures during the recruitment period.

Central to the obtaining of reliable evidence is that good standard care, both surgical and non-surgical, is provided throughout the trial. Where possible, the decisions on the actual method of surgery, when allocated, and non-surgical treatment is left to the clinical judgement of the participating surgeon. Participating surgeons will be advised that they should, however, use surgical interventions and procedures with which they are familiar. This is to avoid learning curve problems. Similarly, physiotherapists are advised that they should use procedures with which they are familiar. The essential components of physiotherapy at each session will be recorded prospectively.

Surgery

For displaced (2 part) surgical neck fractures: surgical interventions with which the surgeon is familiar. These are likely to be plate fixation or intramedullary nailing.

For 3 part (including displaced surgical neck) or 4 part fractures: surgical interventions with which the surgeon is familiar. These are likely to include internal fixation such as nail, plate or other method which preserves the humeral head; or humeral head replacement (hemi-arthroplasty).

Peri- and post-operative management

Peri-operative management including anaesthesia and analgesia, and antibiotic and thromboembolism prophylaxis, dressing policies will follow local guidelines. It is envisaged that similar rehabilitation packages, including mobilisation protocols, should be provided for all interventions. Specifically developed guidelines will be included in the materials for each centre.

Non-surgical intervention (the control group)

Brief recommendations for conservative treatment for trial participants will be included in the materials for each centre. Essentially, these advise that conservatively treated patients will be given sling immobilisation for about 3 weeks or for as long as the treating clinician deems necessary and active early rehabilitation. We will stress the need for competence in conservative methods, including rehabilitation.

Rehabilitation

As far as practical, centres are required to provide written advice on personal care during sling immobilisation to all eligible patients. A generic document has been devised that can be adopted by the centre if required. We will stress that similar access to physiotherapy should be provided for surgical and non-surgical participants. A basic treatment protocol for physiotherapy will be provided. This will emphasise that while the protocol acts as a guide, variation in practice is accepted and anticipated. Electrotherapy other than TENS will be disallowed. We will promote strongly the need to encourage patients to perform home exercises and that they receive information sheets illustrating how to do the exercises.

Outcome measures

The primary outcome measure is the Oxford Shoulder Score (OSS) assessed at 6, 12 and 24 months (Dawson *et al.* 1996). The Oxford Shoulder Score is a condition-specific questionnaire providing a total score based on the patient's subjective assessment of pain and activities of daily living (ADL) impairment. Consistent with recent developments, the range of available scores is 0 (worst) to 48 (best) (Dawson *et al.* 2008). The OSS contains 12 items, each with five categories of response. It has been shown to correlate well with both the professionally-endorsed Constant Score (Constant *et al.* 1987) and the SF36 assessment and to be sensitive to clinical change at six months after surgical intervention (Dawson *et al.* 2001). It has been demonstrated to be consistent, reproducible and valid in a UK population (Dawson *et al.* 1996). This questionnaire will be administered by post for self completion by the trial participant without need for an examination and thus avoids the requirement for follow-up visits to the clinic for assessment. To improve compliance, reminders will be sent and patients will be offered the option of completing the questionnaire via a telephone call. We will also send pre-notification letters and use unconditional incentives; both have been shown to be effective at improving response rates (Edwards *et al.* 2002).

Secondary outcomes are:

- Surgical complications; including shoulder dislocation, failure of implant, proven wound infection (purulent discharge plus positive bacteriology or need for revision due to infection), septicaemia (clinical evidence of systemic infection plus positive blood cultures).
- Early medical complications, i.e. chest infection, confirmed MI or stroke (confirmed by senior clinician), treated DVT, treated pulmonary embolism and other serious event.
- Mortality, subsequent referral for operation or substantive treatment.
- The SF12 and Euroqol (EQ-5D) to collect general health status data (at 6, 12 and 24 months).

Data for economic evaluation

Prospective cost data on trial participants include costs incurred in hospital and subsequently. Thus, time spent in theatre and hospital consumables will be collected. Health utility data will also be obtained from the EQ-5D collected at 3, 6, 12 and 24 months (EuroQol Group 1990). Information for estimating NHS and societal costs will be collected from the trial participants at each follow-up.

We will collect data on the actual procedures done, including anaesthesia, and interventions provided and the experience of operators/care providers (according to grade). We will collect these data for all trial participants.

Data collection

We shall aim to make the trial processes as simple as possible in order to minimise the work entailed at the participating centres. As far as possible, we hope to achieve complete follow-up of all randomised patients.

Baseline data

Basic information including key baseline characteristics will be collected for all potentially eligible patients (i.e. those meeting the trial inclusion criteria) who are found not to be eligible (see Study eligibility form: *Appendix 1*).

Additional data on patient preferences, surgeon's advised treatment and the agreed treatment will be obtained for patients who do not consent to trial participation (see Consent status form: *Appendix 2*).

For consenting patients, we will collect data on ethnicity, education, employment, previous fractures, shoulder dominance, injury mechanism, smoking, diabetes, treatment preference, current health status (EQ- 5D), GP name and surgery and the patient's contact details (see Baseline form: *Appendix 4*).

Description of treatment

Surgical methods

Brief details of the actual surgery and procedures used will be recorded by the surgeon, or assigned deputy, following the operation. Also prescribed rehabilitation.

Non-surgical methods

Brief details of the prescribed non-surgical treatment will be recorded by surgeon, or assigned deputy. Also prescribed rehabilitation.

Collection of hospital outcome data before hospital discharge

Centres will be required to complete data forms detailing:

- Clinical outcomes including surgical complications and early medical complications
- Resource use: the data on hospital costs will be collected using a cost proforma designed for the trial
- Substantive deviations from prescribed treatment and rehabilitation
- Patient destination after hospital discharge.

Long term follow-up

One and two year follow-up forms from centres

Forms to notify mortality or subsequent surgery for completion and return at any time will be made available for the completion by centre staff.

Forms for completion will be sent at 1 and 2 year follow-up – it is likely that all the data for these can be gleaned from the hospital records for these patients.

Follow-up patient questionnaires: 3, 6, 12 and 24 months post trial recruitment

A short questionnaire including the EQ-5D and brief questions on the number of consultations of NHS care providers (GPs, physiotherapists, district nurses etc.), hospitals attendances, use of private healthcare and days lost from work or other normal activities will be sent by the YTU to all participants at three months. See *Appendix 7* for a copy of the 3 month form and covering letter. Reply paid envelopes will be included. Reminders will be sent after 2 and 4 weeks.

Full questionnaires will be sent by the YTU to all participants at six, 12 and 24 months after recruitment. These include the Oxford Shoulder Score, EQ-5D and SF12, all of which are self completion questionnaires. As at 3 months, brief questions on the number of consultations of NHS care providers (GPs, physiotherapists, district nurses etc), hospitals attendances, use of private healthcare and days lost from work or other normal activities will also be requested. See *Appendix 8* for a copy of the 6 month follow up form and covering letter. Reply paid envelopes will be included. Reminders will be sent after 2 and 4 weeks and options for completion of the questionnaires via telephone after 6 weeks. If completed over the telephone, only the Oxford Shoulder Score, EQ-5D and information on hospital readmissions will be requested. An unconditional incentive payment of £5 will be sent for the 12 and 24 month follow-ups (Edwards *et al.* 2002).

X-rays

Copies of all baseline X-rays for all randomised patients will be requested for independent and blinded assessment at the end of study recruitment. X-rays will also be reviewed by local experts on an on-going basis during trial recruitment to ensure the images are of sufficient quality for scrutiny and classification based on the Neer classification system by an independent panel of experts at the end of recruitment.

Proposed sample size

The primary outcome for the trial is differences in patients' subjective assessments of pain and activities of daily living (ADL) as measured by the Oxford Shoulder Score (OSS). For surgery to be worthwhile, it needs to demonstrate greater improvements in patient's subjective assessments of pain and ADL than those for conservative treatment to justify both its increased costs and the exposure to the hazards of surgery. In an observational study conducted by one of us (AR) it was found that those patients who had surgery had a 5 point differential improvement in the OSS compared with those patients treated conservatively. Given a SD of 12 this equates to an effect size of 0.42. We propose, therefore, to design the study to observe an effect size of 0.4 at 80% power using 5% significance level, which would require approximately 200 participants. After allowing for drop-outs of 20%, we propose to recruit and randomise 250 patients (125 surgery and 125 controls). Our estimate of 20% loss from the RCT is purposefully pessimistic for sample size calculations.

Recruitment rate

We anticipate that recruitment for this trial will be potentially challenging. Therefore, we have set very conservative recruitment targets. Our recruitment period is 18 months. We aim to recruit between 18–20 centres. Each centre will be expected to recruit only one participant per month, though encouraged to aim higher than this. We estimate across 18 centres there will be 6066 patients (6000 is used as a working figure here) with a proximal humeral fracture over the 18 months of recruitment. Of these, 2391 (thus, 2400 is used as a working figure below) will have the fracture types suitable for inclusion into the RCT. To achieve our sample size we need to recruit only 11% of these patients.

Loss to follow-up

We anticipate that the main reason for loss to follow-up will be mortality. We will follow up patients assiduously using postal questionnaires and in the event of non-response we will contact their GP to ascertain whether or not it is appropriate to contact the patient and, if so, their address.

Statistical analysis

All of the analyses will use the intention to treat principle. Consequently, any patients who cross over from either study arm will be analysed as per their randomisation status. The primary outcome is the Oxford Shoulder Score (OSS). The difference between the two treatment groups will be compared over all follow-up assessments (i.e. 6, 12 and 24 months) using a repeated measures model. The model will include terms for treatment, follow-up time, and also adjust for type of fracture, age and gender (as older people and women are more likely to sustain these fractures). Because participants are clustered by surgical centre there is a theoretical possibility that there may be a 'surgeon' effect. We will therefore repeat the primary analysis using appropriate statistical techniques (robust standard errors) to account for the clustering of patients within surgeon. The anonymity of individual surgeons and centres will be preserved for all analyses and there will be no presentation or comparisons of the treatment results from individual centres or surgeons. Subgroup analyses based on the Neer classification system are planned to assess the effectiveness of treatment for the different fracture groups (2 part surgical neck; 3 part including surgical neck and 4 part fractures; fractures not meeting the Neer classification displacement criteria). The secondary outcomes will be summarised for each treatment group.

Frequency of analyses

We anticipate that there will be a single analysis at the end of the study. However, before the study starts we will establish an independent Trial Steering Committee (TSC) and a Data Monitoring and Ethics Committee (DMEC). The chairperson of the TSC will make the decision in conjunction with the chair of the DMEC about the need and the number of interim analysis.

Assessment of study recruitment and applicability

We will report the numbers of and reasons for ineligible adults (aged 16 or above) with the study fractures, we will also report the numbers of and reasons for the non-inclusion of potentially eligible patients. We will compare the baseline characteristics, patient preferences with those of randomised patients.

Economic evaluation

An economic analysis will be taken from the perspective of the UK National Health Service and Social Services. The horizon for the baseline analysis will be two years. However, we will model any potential benefits forward to 5 years and an average lifetime in a sensitivity analysis.

Health benefits for the economic analysis will be measured in terms of quality-adjusted life-years (QALYs). Health utility values for individuals with displaced proximal humeral fractures will be estimated using the Euroqol (EQ-5D) questionnaire. QALYs will be calculated for each patient using the area under the curve defined by her/his EQ-5D scores over the two-year follow-up period and adjusted by the Kaplan Meier estimates of patients' survival over the same period of time. Given the horizon for the analysis is longer than a year a discount rate of 1.5% will be applied to health benefits (NICE 2001).

Resource use, and clinical data will be collected for all trial participants. Information regarding total volume of resources used in the treatment (conservative/surgical) and rehabilitation procedures will be recorded for each patient. Unit cost will then be applied to estimate the total cost per patient. To account for the censored nature of cost data, the Lin method will be used to estimate the mean average total cost per treatment arm (Lin *et al.* 1997). Non parametric bootstrapping techniques will be used to estimate 95% confidence intervals for the mean differential cost between conservative and surgical treatment (Efron *et al.* 1993). Total cost will be discounted using a 6% annual discount rate (NICE 2001).

Health benefits and mean average total costs associated with each of the trial arms will be combined in a cost-utility analysis, incremental costs per quality ratios will be computed comparing the conservative and surgical treatment interventions for adult patients with a displaced proximal humeral fracture. Multilevel modelling will be used to explore potential variations in treatment effect and costs between health professionals (Roberts 1999).

Project timetable and milestones

Currently (December 2007), the anticipated start for full-trial recruitment is 1st October 2008. As well as MREC approval, all trial materials and processes will have been established and piloted where practical and appropriate at Middlesbrough, the lead centre. Applications for LREC approval will also have been submitted, and may have been approved, for several centres. Staff training for these centres will also have been initiated. The 10 months target date after recruitment for the feasibility study means that a robust approach is needed for getting centres started and recruiting as soon as possible.

Trial timeline from start of full-trial recruitment

0 months

- Start of recruitment for RCT: duration 18 months (assuming satisfactory outcome of the 10 month feasibility study)
- Start of data collection for baseline, treatment, early complications and economic outcomes

3 months

- Start of 3 month follow-up for RCT participants (reminders 2 & 4 weeks).

6 months

- Start of 6 month follow-up for RCT participants (reminders 2 & 4 weeks).

10 months

- Cut-off for analysis of feasibility study after 10 months of recruitment.
- Analysis and presentation of feasibility study.

12 months

- End of recruitment of feasibility study.
- Start of 12 month follow-up for RCT (reminders 2 & 4 weeks).

18 months

- End of all recruitment (RCT).

21 months

- End of 3 months RCT follow-up (reminders 2 & 4 weeks).

24 months

- End of 6 months RCT follow-up (reminders 2 & 4 weeks).
- Start 24 months RCT follow-up (reminders 2 & 4 weeks).

30 months

- End of 12 months RCT follow-up (reminders 2 & 4 weeks).

42 months

- End 24 months RCT follow-up (reminders 2 & 4 weeks).
- Analyses.
- Preparation of HTA report.
- Preparation of study publications.

Complete project: 48 months inclusive

If recruitment is stopped at the conclusion of the 12 months feasibility study, the study will complete at 42 months.

Dissemination of trial findings

We shall disseminate our findings through relevant local, national and international conferences and peer-reviewed publications. Reflecting the collaborative basis of this research, all active contributors will be named and credited in the main report.

Trial management

The day to day management of the project is the responsibility of the Trial Management Group:

- Clinical co-ordination: Amar Rangan (Chief Investigator)
- Trial management: Stephen Brealey (Trial Manager, University of York) and Laura Dennis (Trial Co-ordinator, University of Teesside)
- Methodological support: Helen Handoll (University of Teesside), David Torgerson (University of York).

Clinical queries should be directed to any of the listed clinical co-ordinators, and any other issues may be discussed with any member of the Trial Management Group.

The trial co-ordinating centre is York Trials Unit. Specifically assigned to the ProFHER trial are a Statistician (Mrs Gill Worthy, replaced by Dr Catherine Hewitt (April 2009)), a Health Economist (Dr Jo Dumville, replaced by Miss Ling-Hsiang Chuang (May 2009)), Data Managers (Mr Ben Cross and Mrs Valerie Wadsworth), and a Trial Secretary (Mrs Sarah Gardner).

Research governance

The University of Teesside, of which the Chief Investigator holds an honorary lecturership position, is the sponsor.

Independent supervision

A Trial Steering Committee (TSC) will be established upon acceptance by the HTA of our nominees for an independent chairperson and two other independent members. Other members will be those key to trial management and function. Accrual, whole population and 'housekeeping' statistics will be calculated and disseminated to members on a monthly basis to monitor the progress of the RCT. The TSC will meet at the discretion of the independent chairperson, however, we anticipate that it will meet approximately a total of nine times throughout the study.

A separate DMEC will be established that is independent of the applicants and the TSC. The names of three nominees, a trial statistician, a shoulder specialist and a consultant rheumatologist, who have accepted our invitation to fulfil this role have been submitted to the HTA. We anticipate that the DMEC will meet a total of six times, approximately every 6 months, or at any other time if any issues were to arise that required urgent attention. An analysis of major complications will be provided for the DMEC every three months following the commencement of the recruitment phase of the trial.

Trial funding

The PROFHER trial is funded by the NHS R&D Health Technology Assessment Programme (HTA).

Reimbursement for centres

Centres will be advised to consider employing dedicated clinical staff, such as a research nurse, to act as the contact person for recruitment and data collection for each centre. This could involve increasing someone's hours.

The HTA has agreed to a reimbursement scheme that recognises the additional efforts and resources required for recruitment into this trial. The following payments have been proposed:

- (a) Receipt of eligibility form for non-eligible patients (see *Appendix 1*): £10
- (b) Receipt of eligibility and consent status forms where patient consent has been sought: £30
- (c) For each randomised patient. Provisional scheme: upon receipt of baseline materials and description of treatment received (£500); receipt of pre-discharge information and copies of X- rays (£250); all remaining information at 1 and 2 year follow-up (£250).

The following activities are covered by these reimbursement payments

- Distribution of trial materials in appropriate locations
- Setting up the trial processes
- Screening of patients' notes for eligibility
- Contacting patients or alerting appropriate clinicians for this activity
- Discussing the trial with patients and providing documentation for consideration
- Completion of trial eligibility and consent status forms
- Obtaining consent
- Randomisation of individual patients and completion of baseline form
- Completion of forms for treatment, costs and in-hospital outcome and complications for randomised patients.
- Completion of forms for randomised patients at one and two years.
- Copying and carriage of baseline X-rays for independent review.

Health service costs

We anticipate that there will be no excess treatment costs overall. Whilst this is likely to be the case in the majority of centres, this may not necessarily apply to certain hospitals, even in the context of a potentially low recruitment rate, that have a below average rate of surgery. It will be up to the individual providers to consider this aspect in giving their approval for the trial. Some check at randomisation to prevent excess surgery (especially in small centres) is being considered at present.

References

- Brorson S, Bagger J, Sylvest A, Højbjerg A. Improved interobserver variation after training of doctors in the Neer system: a randomised trial. *Journal of Bone & Joint Surgery – British Volume* 2002;**84**(7):950–954.
- Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clinical Orthopaedics and Related Research* 1987;**214**:160–164.
- Court-Brown CM, Garg A, McQueen MM. The epidemiology of proximal humeral fractures. *Acta Orthopaedica Scandinavica* 2001;**72**(4):365–371.
- Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. *Journal of Bone & Joint Surgery – British Volume* 1996;**78**(4):593–600.
- Dawson J, Hill G, Fitzpatrick R, Carr A. The benefits of using patient-based methods of assessment. *Journal of Bone & Joint Surgery – British Volume* 2001;**83**:877–882.
- Dawson J, Rogers K, Fitzpatrick R, Carr A. The Oxford shoulder score revisited. *Archives of Orthopaedic Trauma Surgery* 2008 (Epub). DOI: 10.1007/s00402-007-0549-7.
- Edwards P, Roberts I, Clarke M, DiGuseppi C, Pratap S, Wentz R, Kwan I: Increasing response rates to postal questionnaires: systematic review. *BMJ* 2002;**324**:1–9.
- Efron B, Tibshirani RY. *An Introduction to the Bootstrap*. 1993. New York: Chapman & Hall.

EuroQol Group: EuroQol: a new facility for the measurement of health related quality of life. *Health Policy* 1990;**16**:199–208.

Handoll HHG, Gibson JNA, Madhok R. Interventions for treating proximal humeral fractures in adults. *Cochrane Database of Systematic Reviews* 2003; Issue 4. (minor update: Issue 2, 2007).

Handoll HHG, Madhok R. From evidence to best practice in the management of fractures of the distal radius in adults: working towards a research agenda. *BMC Musculoskeletal Disorders* 2003;**4**(27).

Hodgson S. Proximal humerus fracture rehabilitation. *Clinical Orthopaedics & Related Research* 2006;**442**:131–138.

Lin DY, Feuer EJ, Etzioni R, *et al.* Estimating medical costs from incomplete follow-up data. *Biometrics* 1997;**53**:419–434.

Misra A, Kapur R, Maffulli N. Complex proximal humeral fractures in adults – a systematic review of management. *Injury* 2001; **32**(5):363–372.

Moher D, Schulz KF, Altman DG, for the CONSORT group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;**357**:1191–1194.

Neer CS. Displaced proximal humeral fractures. Classification and evaluation. *Journal of Bone and Joint Surgery American Volume* 1970;**52**(6):1077–1089.

Neer CS. Four-segment classification of proximal humeral fractures: purpose and reliable use. *Journal of Shoulder and Elbow Surgery* 2002;**11**(4): 389–400.

NICE. *Technical Guidance for Manufacturers and Sponsors on Making Submissions to Technology Appraisal*. 2001. London, National Institute for Clinical Excellence.

Palvanen M, Kannus P, Niemi S, Parkkari J. Epidemiology of proximal humeral fractures. *Clinical Orthopaedics and Related Research* 2006;**442**:87–92.

Roberts C. The implications of variations in outcome between health professionals for the design and analysis of randomised controlled trials. *Statistics in Medicine* 1999;**18**:2605–2615.

Solomon MJ, Mcleod RS. Should we be performing more randomised controlled trials evaluating surgical operations? *Surgery* 1995;**118**(3):459–467.

Appendix 4 Changes to the protocol

This appendix describes changes to the protocol; the final protocol is presented in *Appendix 3*.

Recruitment period

Cancellation of the feasibility study

In the protocol, as well as stipulating a recruitment period of 18 months with contributions from at least '18 NHS trauma centres', there was an assessment point after 10 months of recruitment (1 August 2009) to determine whether or not recruitment into the RCT was sufficient (set at 88 patients) to justify its continuation after 12 months. It soon became clear that we would not make this target (31 participants were recruited after 10 months). With the support of the TSC, however, we successfully put forward a case for cancelling this feasibility study and associated stopping rule for cessation of recruitment at 12 months. Our case was based on evidence that patient consent was higher than expected – at that stage, of 32 patients approached, 20 had consented – and that the serious delays in acquiring R&D approval for setting up sites to recruit had significantly restricted the potential of the trial to meet this preliminary target.

Extension

Subsequently, we applied for a funding and time extension (for 18 months). Our proposal included a revised forecast based on doubling the original number of centres but reducing the anticipated recruitment to one patient every 2 months per centre. Additionally, we tightened up our approach towards potential centres by establishing criteria for whether or not to expend resources, primarily trial manager time, on setting up a new centre. We remain grateful to our funders for 'keeping faith' with this trial at this critical time and for the practical support of the independent members of the TSC and DMEC.

Terminology

Reconfiguration of trauma care in the UK subsequent to the end of the trial resulted in designated 'trauma centres'. We thus changed the description of 'trauma centres' to hospitals, to avoid confusion.

Randomisation

Originally, we planned to start with simple randomisation, stratified by the presence or not of a tuberosity fracture and with a prespecified block size, followed by a prespecified period switch to randomisation using a computer-generated minimisation programme. The intended minimisation factors were centre and fractures involving either tuberosity. From the start, random block sizes were selected to prevent predictability of the sequence. Minimisation was also not implemented because, with fewer patients being recruited at centres than originally anticipated, the rationale for minimisation by centre became questionable. There was a potential for predictability of the sequence, and the low recruitment rate and lower rate of recruitment at individual centres meant that there were not the anticipated logistical and cost problems stemming from an excess of patients needing an operation at individual centres.

Study inclusion/exclusion criteria

We decided, with approval from the TSC, that the displacement of the surgical neck did not have to meet the exact Neer displacement criteria (1 cm and/or 45° angulation of displaced parts) for inclusion of a fracture in the trial. This is because of the arbitrary thresholds for defining displacement and the advice

from the TSC to be more pragmatic about the exact displacement. Instead, there needed to be surgical neck 'involvement'; however, there was emphasis on displacement of the surgical neck throughout the trial documentation and this was implicit also in the continual reference made to Neer two- and three-part fractures involving the surgical neck.

Interventions

Non-surgical treatment

As indicated in the final version of our protocol, we decided against the development and provision of consensus guidelines for non-surgical treatment. Instead, we emphasised the provision of good standard care and indicated both verbally and in the trial manual that we anticipated that initial care would comprise sling immobilisation for about 3 weeks or for as long as the treating clinician deemed necessary and active early rehabilitation.

Home exercises

Throughout the trial we planned to promote home exercises. However, as stated in the final version of our protocol, we realised that our original intention to produce generic information leaflets illustrating exercises for home use by patients was misguided given the expected variation in patient ability and changes to requirements over time. Instead, we indicated that physiotherapists would either provide these or access standard web-based facilities to generate 'bespoke' exercise sheets. However, feedback from the lead physiotherapy contact for the trial (LG) after discussions with the physiotherapists acting as the named contacts for physiotherapy at the participating hospitals indicated that advice on this aspect was often adjusted to accommodate home exercises that were predominantly based on daily functional tasks, for which written instructions were considered unhelpful. Instead, the emphasis was on the monitoring and reinforcement of these and related functional activities at subsequent physiotherapy.

Outcomes

The secondary outcomes listed in the protocol were as follows:

- surgical complications, including shoulder dislocation, failure of implant, proven wound infection (purulent discharge plus positive bacteriology or need for revision because of infection) and septicæmia (clinical evidence of systemic infection plus positive blood cultures)
- early medical complications, that is, chest infection, confirmed MI or stroke (confirmed by senior clinician), treated DVT, treated pulmonary embolism and other serious event
- mortality, subsequent referral for operation or substantive treatment
- general health status assessed using the SF-12 and EQ-5D (at 6, 12 and 24 months).

The following adjustments were made to the secondary outcomes, which were reordered to give higher precedence to health status data as reported by the patient:

- it was decided to present SF-12 data as PCS and MCS scores
- EQ-5D data are presented as part of the economic evaluation
- it was clarified that other shoulder fracture-related complications would be collected, thus removing the unintended emphasis on complications solely related to surgery
- reflecting the wording on the data collection forms, 'subsequent referral for operation or substantive treatment' was modified to 'secondary surgery to the shoulder or increased/new shoulder-related therapy'
- reflecting the data collection process, it was clarified that we were focusing on medical complications experienced during the inpatient stay
- mortality was presented as a separate outcome.

Data for the economic evaluation

For the collection of data on hospital consumables and costs it was clarified that these were data relating to surgery, the inpatient stay and hospital physiotherapy.

Decisions on the selection of the sources of data, from those reported in the hospital forms and patient questionnaires, for the base-case analysis were not stipulated in the protocol but were prespecified in the economic analysis plan. These were based on the perceived reliability and comprehensiveness, in terms of cost information and data completeness, of the source data and relationships with similar data categories.

As planned, data for estimating societal costs were collected from the trial participants at each follow-up.

Radiological assessment to determine study eligibility

The initial intention to require hospitals to use the full shoulder trauma series (thus, three perpendicular views)⁵⁹ was adjusted, after consultation with the independent members of the TSC, to promote the need to adhere to local guidelines (always two views). This was more in keeping with the pragmatic intention of the ProFHER trial and addressed the need to seek local approvals for additional views, which was causing delays in the setting up of sites.

Ongoing review of baseline radiographs

For practical reasons, including the greater than anticipated difficulties in obtaining copies of images that could be accessed independently of the PACS at various individual hospitals, the original intention of carrying out ongoing monitoring and quality review of baseline radiographs during trial recruitment was revised. The audit of all 250 radiograph sets was completed after the end of recruitment. Additionally, the Chief Investigator was assisted in the review of the quality of the copies of the radiographic images for each trial participant by one orthopaedic surgeon (Mr Jaime Candal-Couto) instead of the two originally planned. In the final stage of the quality audit, the independent orthopaedic surgeon member of the TSC (Professor Andy Carr) acted as a third rater for resolving disagreements.

Classification of the radiographs

Independent classification of the radiographs according to the Neer classification system was performed, after training, by two orthopaedic specialist shoulder surgeons experienced in assessing and operating on proximal humeral fractures. Both surgeons, who had previous experience with the Neer classification system, were not involved in the trial but were representative of surgeons who would have recruited patients in a participating site. This replaced our original intention for classification to be carried out by three 'experts' forming an independent panel of musculoskeletal radiologists or orthopaedic surgeons who had experience with the Neer classification. To address the known problem of poor inter-rater agreement we developed a proforma to collect data and undertook a pilot classification and training exercise with two senior registrars in the final year of their training at the lead site (James Cook University Hospital) using a sample of radiographs. Based on the findings of this pilot, the two independent surgeons underwent systematic training before performing a two-staged and fully documented process, involving independent classification of the radiographs and discussion to reach a final verdict.

Clinical effectiveness analysis

Subgroup analyses

We added in an a priori subgroup analysis based on age (< 65 years, ≥ 65 years) and revised our subgroup analysis of fracture classification. The latter was based on recorded tuberosity involvement at baseline (yes/no) and replaced our original intention to split the fractures into three groups (two-part fractures involving the surgical neck; three-part fractures involving the surgical neck and four-part fractures; and fractures not meeting the Neer classification displacement criteria). We revised this for consistency of

reporting (tuberosity involvement being reported on the study eligibility form for all eligible patients) and to keep to the original subgrouping in terms of study design (two-part fractures vs. three- and four-part fractures). We conducted a supplementary sensitivity analysis comparing Neer one- and two-part fractures with Neer three- and four-part fractures based on the results of the independent classification of baseline fractures.

Cluster analysis by surgeon

In the protocol we indicated that we would repeat the primary analysis using appropriate statistical techniques to account for the clustering of patients within surgeons, thus testing for a potential 'surgeon' effect. As surgeons at some hospitals operated on very few patients, we conducted the analysis at a centre level; this quantified the performance of the surgical team(s) at the hospital rather than individual surgeons.

Additional sensitivity analysis

After recruitment to the trial had finished, an imbalance in smoking status was observed. Baseline imbalance in itself is not considered an appropriate reason to include a baseline measure as a covariate. However, smoking is associated with a number of complications including prolonged fracture and wound healing (including surgical site infection), impaired new bone formation and development of osteoporosis.^{105–107} Hence, as endorsed by the DMEC, a sensitivity analysis was undertaken in which smoking status was included in the primary outcome model to explore whether or not the conclusions drawn from the primary analysis are robust.

Economic analysis

Missing data and censoring issues

We used multiple imputation^{56,57} rather than Lin's methods,⁵⁸ as proposed in the protocol, for dealing with missing data and censoring issues relating to cost data. This is because, although dealing with censoring issues, Lin's methods^{58,108} do not take 'inter-missing' data, which are missing observations occurring before any last observation, into account. Consider, for instance, a participant with observations at 3 and 12 months' follow-up only. In this case, apart from censoring after 12 months, the missing observations at 6 months are called inter-missing. Furthermore, Lin's methods do not allow for exploration after the follow-up of the trial; this became relevant because of a subsequent decision to perform a 5-year long-term follow-up. Lastly, the multilevel modelling approach used for the statistical analysis could not be implemented using Lin's methods.

Despite the fact that we did not expect any substantial differences between estimates based on multiple imputation and those based on Lin's methods, we considered that there were several advantages of multiple imputation over Lin's methods. First, both issues of censoring and inter-missing could be addressed. Second, both multiple imputation and multilevel analysis allow for extrapolation. Therefore, multiple imputation using chained equations¹⁰⁹ was adopted to address the missing data in the final economic analysis.

Discount rates

We used updated discounting rates according to the latest NICE guidance,⁵⁵ which differed from the rates based on 2001 figures stated in the protocol (1.5% for health outcomes and 6% for total costs). Discount rates were set at 3.5% for both costs and benefits.

Long-term sensitivity analysis

In the protocol we indicated that we would model any potential benefits forward to 5 years (from 2 years) and an average lifetime in a sensitivity analysis. Because of the subsequent set-up of a long-term follow-up study aiming to collect patient outcomes up to 5 years, we decided that, in the current economic analysis, we would not conduct any exploration beyond the length of the 2-year follow-up. Instead, we intend to produce a separate paper discussing the forecast and observed outcomes once the long-term follow-up study is completed.

Appendix 5 Study eligibility form

Version 5.0 (19/05/09)

ProFHER TRIAL: STUDY ELIGIBILITY FORM

Participant identification number:

This form is for orthopaedic specialists to complete when assessing patient's eligibility to enter the trial.

Please complete this form for any patient aged 16 or above, presenting within three weeks of their injury with a radiologically confirmed displaced fracture of the proximal humerus with involvement of the surgical neck. This includes all Neer 2, 3 part fractures involving the surgical neck and 4 part fractures.

Name of Orthopaedic Specialist:

Patient's Hospital Number:

Date today:

 / /

Day Month Year

Date of injury to shoulder:

 / /

Day Month Year

Patient's sex:

(please cross one box)

Male

☐

Female

☐

Date of birth:

 / /

Day Month Year

Affected shoulder:

(please cross one box)

Left

☐

Right

☐

The displaced fracture **MUST** involve the surgical neck for inclusion in the trial. In addition, does the fracture involve:

(please cross all boxes that apply)

Greater tuberosity

☐

Lesser tuberosity

☐

What series of X-rays were used?
(please cross all boxes that apply)

Anteroposterior view

☐

Axillary (or modified axillary) view

☐

Scapular Y-Lateral

☐

Exclusion criteria: (please cross all boxes that apply)

a. Associated dislocation of the injured shoulder joint:

☐

f. Pathological fracture (but not osteoporotic):

☐

b. Open fracture:

☐

g. Terminal illness:

☐

c. Other upper limb fractures:

☐

h. Patient not resident in trauma-centre catchment area:

☐

d. Clear indication for surgery such as severe soft tissue compromise requiring surgery or emergency treatment (e.g. neurovascular injury or dysfunction):

☐

i. Mentally incompetent (unable to understand the trial or instructions for rehabilitation):

☐

e. Co-morbidities that preclude surgery or anaesthesia:

☐

j. Other reason to exclude the patient:

☐

If you crossed the box for j. please record your reason in the box below:

Is the patient eligible? (please cross one box)

Yes

☐

No

☐

If the patient is eligible then the designated person(s) in your hospital should obtain patient consent.

If the patient is not eligible then what treatment would you advise for this patient?
(please cross one box)

Surgery

☐

Not Surgery

☐

Thank you for completing this form. Could you now please give this form to the designated person so that the form can be checked for completeness and posted to York Trials Unit.



PROximal Fracture of Humerus: Evaluation by Randomisation (ProFHER) Trial
A multi-centre randomised controlled trial funded by NHS R&D Health Technology Assessment Programme
(International Standardised Randomised Controlled Trial Number 50850043)

0177445626

Appendix 6 Sample population estimates

The following section, reporting work undertaken in 2007, explains the basis for our estimates of the number of sites and recruitment period needed to achieve the recruitment target prior to setting up of the trial.

A key reference by Court-Brown *et al.*⁷ supplies the incidence of proximal humeral fractures in a known population in Edinburgh and a breakdown by fracture type (displaced 51%; two part involving the surgical neck 28%; three part involving the greater tuberosity and surgical neck 9%; and four part without fracture dislocation 2%). Apart from this study there was a lack of reliable data on the annual incidence of people being treated for proximal humeral fractures in other hospitals in the UK. Indeed, the lack of information on patient numbers found early on from putative principal investigators continued subsequently when attempting to obtain estimates when recruiting centres into the trial. There was also a lack of catchment population data for individual hospitals. Using publicly available HES data appeared to offer a solution.

The 2004–5 HES data showed that there were 9378 FCEs relating to the care of proximal humeral fractures out of a total of 13,706,765 FCEs in that period (see *Appendix 1*). Assuming that a similar proportion of proximal humeral fractures were treated each year compared with the total number of FCEs in each of the six original centres (principal investigators were grant applicants), the following table was constructed. The first estimate was based on all FCEs and the second on actual admissions (there were 7600 listed at the time according to HES 2004–5 data).

Centre ^a	Hospital provider	All FCEs	Estimate 1: proximal humerus fracture FCEs ^b	Estimate 1: proximal humerus fracture FCEs ^c
Coventry and Warwick	University Hospitals Coventry and Warwickshire NHS Trust	105,119	72	58
Leicester	United Hospitals of Leicester NHS Trust	237,813	162	132
Newcastle	Newcastle upon Tyne Hospitals NHS Foundation Trust	175,356	119	96
Nottingham	Queen's Medical Centre	119,323	82	66
Stoke-on-Trent	North Staffordshire Hospital NHS Trust	152,638	104	84
Teesside	South Tees Hospitals NHS Foundation Trust	146,356	100	81

a All six centres participated in the trial. Three centre names were changed subsequently: 'Coventry and Warwick' was referred to as 'Coventry'; 'Stoke-on-Trent' was referred to as 'North Staffs'; and 'Teesside' was referred to as 'James Cook'.

b Based on 9378 (number of FCEs relating to fracture of the upper end of the humerus)/13,706,765 (total number of FCEs) × FCEs at hospital provider.

c Estimate 1 × 0.81 to reflect the fact that 7600 (81%) of the 9378 FCEs relating to fracture of the upper end of the humerus were admission FCEs.

The following estimates were based on 'estimate 1', thus on all FCEs. We assumed that all hospital admissions would be displaced fractures and that these would represent 51% of the total proximal humeral fractures⁷ that would present to the hospital. The 'exclusion' of undisplaced fractures from the HES data related to the perception that patients with these fractures would not be admitted. Therefore, in the six hospitals, 639 fractures would be displaced proximal humerus fractures and there would be 1253 fractures in total, of which 350 would be two-part fractures involving the surgical neck and 138 would be three- or four-part fractures.

Based on the average for five hospitals (excluding Leicester, an outlier), 382 displaced fractures were added for four extra hospitals; it was considered that the addition of four extra hospitals would be enough at that time. Therefore, for nine hospitals, 1021 fractures would be displaced fractures and there would be 2044 fractures in total, of which 572 would be two-part fractures involving the surgical neck and 225 would be three- or four-part fractures.

The uncertainty in these annual estimates was emphasised. Although there was good correspondence between the prospective data received from the principal investigator in Teesside (100 fractures annually in the orthopaedic department) and estimate 1 for Teesside, a much lower estimate of 100 was received from the principal investigator at Leicester than estimate 1 for this centre ($n = 162$).

In 2007 we took the decision to double the number of centres to 18 and to recruit for a period of 18 months. We therefore estimated that there would be 6132 patients from these centres with a proximal humeral fracture, of whom 2391 would have the fracture types suitable for inclusion into the trial. This would mean that only 11% of the patients identified would need to be enrolled into the study to meet our recruitment target of 250 patients, which we judged to be feasible.

Appendix 7 Table of amendments sent to the Multicentre Research Ethics Committee

Multicentre Research Ethics Committee approval was obtained from the York Research Ethics Committee (reference number 08/H1311/12) – date of favourable ethical opinion 11 March 2008 (protocol version 1, 13 December 2007).

Amendment	Date	Description of main items in the request for approval [including version (v) of protocol if revised]
Minor 1	8 April 2008	Changes, mainly relating to wording, to several forms: study eligibility form changed to clarify that all Neer two- and three-part fractures involving the surgical neck and four-part fractures should be considered, and 'terminal illness' made a separate exclusion criterion
Minor 2	10 June 2008	Protocol v2 (4 June 2008) Changes to the protocol and study eligibility form. Additional exclusion criterion: associated dislocation of the injured shoulder joint. Adjustment of the protocol to reflect the revised scoring scheme for the OSS (high = best health) and the plan for an interim review of the radiographs early in the recruitment period to monitor the eligibility of patients entering the trial
Minor 3	22 July 2008	Protocol v3 (30 June 2008) Changes to several forms (mainly for ease of data collection and processing), letters and the protocol. Submission of the sling immobilisation leaflet. Copies also provided of all letters for contacting patients at follow-up. Key changes to the protocol were clarification of the consensus guidelines for non-surgical treatment; addition of the information leaflet on care during sling immobilisation; and clarification of the consensus guidelines and data collection for physiotherapy treatment, including advice regarding home exercise
01 Substantial	28 August 2008	Protocol v4 (22 August 2008) Measurement of EQ-5D at baseline, reflected in changes in the baseline information form and protocol
02 Substantial	6 November 2008	Protocol v5 (6 November 2008) Randomisation sequence changed from 'Stratified randomisation will be performed by displaced surgical neck fractures without versus with involvement of either tuberosity' to 'After an initial, prespecified period of randomisation, stratified by the presence or not of a tuberosity fracture and with a prespecified block size, randomisation will then be performed using a computer-generated minimisation programme. The minimisation factors will be fractures involving either tuberosity and centre'. Change informed by Professor Martin Bland and endorsed by the trial statistician (Dr Brian Faragher), an independent member of the TSC. Note: minimisation was never employed
03 Substantial	6 April 2009	Protocol v6 (6 April 2009) Changes to the protocol to clarify that a minimum of two radiographic views/projections are required for the assessment of trial eligibility and to amend the interim assessment of radiographs used to determine study eligibility of trial participants to include the assessment of the quality of all radiographs on an ongoing basis during the recruitment period

Amendment	Date	Description of main items in the request for approval [including version (v) of protocol if revised]
Minor 4	19 May 2009	Protocol v7 (19 May 2009)
		Changes to the study eligibility form and the protocol to clarify that all Neer two-part surgical neck fractures, three-part fractures (including the surgical neck) and four-part fractures of the proximal humerus were eligible for inclusion
04 Substantial	1 July 2009	Trial poster to publicise the trial for display on the wards or in fracture clinics as a reminder about the trial and who to contact when a potentially eligible patient is identified. There was no change to the study screening or recruitment processes
Minor 5	23 September 2009	Consent status form amended to allow for clinicians to have continued uncertainty over what treatment they would advise for a non-consenting patient
05 Substantial	28 July 2010	Three new questions included on the last page of the 24-month patient questionnaire. These asked (a) how the participant's shoulder was compared with 1 year ago (transition question for assessing minimally important change in the primary outcome: OSS); (b) the participant's treatment preferences at the end of the trial; and (c) whether or not the participant would like to be informed about the results of the study
06 Substantial	9 September 2010	Protocol for long term follow-up study v1 (9 September 2010)
		Extended follow-up of trial participants, with questionnaires sent at 3, 4 and 5 years for key outcomes. Support confirmed from the trial sponsor, independent members, including the service users, of the TSC and the members of the trial's DMEC
07 Substantial	14 July 2011	Protocol for long-term follow-up study v2 (30 June 2011)
		For sending participants approached to take part in the follow-up study reminder letters at 2 and, if necessary, 4 weeks to enhance the consent rate
Minor 6	4 September 2012	Protocol for long-term follow-up study v3 (9 August 2012)
		Mortality dropped as a secondary outcome for the long-term follow-up study

Appendix 8 The Trial Management Group and membership of the Trial Steering Committee and Data Monitoring and Ethics Committee

Trial management

The day-to-day management of the project was the responsibility of the TMG:

- clinical co-ordination: Amar Rangan (chief investigator) and Lorna Goodchild (specialist physiotherapist)
- trial management: Stephen Brealey (trial manager, University of York) and Laura Jefferson, née Dennis [Trial Co-ordinator, Teesside University (up to September 2011)]
- methodological support: Helen Handoll (Teesside University) and David Torgerson (University of York).

The trial co-ordinating centre was the YU. Specifically assigned to the ProFHER trial were a statistician (Mrs Gill Worthy, replaced by Dr Catherine Hewitt April 2009, Mr Arthur Kang'ombe January 2011 and Ms Ada Keding October 2012); a health economist (Dr Jo Dumville, replaced by Miss Ling-Hsiang Chuang May 2009 and Mrs Belen Corbacho January 2012), data managers (Mr Ben Cross and Mrs Valerie Wadsworth) and a trial secretary (Mrs Sarah Gardner, replaced by Mrs Sue Collins March 2011 and Mrs Angela Rogers May 2012).

Trial Steering Committee

Independent members:

- Professor Sarah Purdy (chairperson and independent member), Professor of Primary Care, University of Bristol
- Professor Andrew Carr (independent member), Nuffield Professor of Orthopaedic Surgery, University of Oxford
- Professor Brian Faragher (independent member), Professor in Medical Statistics, Liverpool School of Tropical Medicine
- patient representatives: Mrs Dorothy Anderson and Mrs Margaret Newlands
- sponsor representative: Dr Alasdair MacSween, Principal Lecturer in Research Governance, Teesside University.

Data Monitoring and Ethics Committee

Independent members:

- Professor Marion Campbell (chairperson), Director of the Health Services Research Unit, University of Aberdeen
- Professor Roger Francis, Consultant Physician, Freeman Hospital, Newcastle upon Tyne
- Mr David Limb, Consultant Orthopaedic Surgeon, St James' University Hospital, Leeds.

Appendix 9 Patient information sheet

Hospital logo

Version 3.0 (08/04/2008)

Patient Information Sheet

PROximal Fractures of the Humerus: Evaluation by Randomisation The PROFHER trial

A UK multi-centre study of surgery for displaced fractures at the top end of the upper arm bone

(Reference number: 08/H1311/12)

We would like to invite you to take part in this research study. Before you decide, it is important for you to understand why this research is being done and what it would involve for you. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

Thank you for reading this

What is the purpose of the study?

Your doctor will have told you that the top end of your upper arm bone (humerus) has broken (fractured) near the shoulder joint. These fractures are common injuries, particularly in older people. The less serious fractures are treated by supporting the arm in a sling to help the bone to heal enough to allow movement. This is followed by exercises.

About half of these fractures are more serious because the parts of the bone are displaced from each other. Such fractures are more difficult to treat and the surgeon may consider that an operation would be helpful. Surgery often involves putting the fractured parts together again and fixing them in place with various devices such as nails or plates. Sometimes the top end or head of the humerus is replaced by an artificial joint. However, for the majority of serious fractures surgeons are not sure whether an operation is better than supporting the arm in a sling and leaving the bone to heal. This is because there is no strong evidence that surgery gives a better result.

The aim of this UK multi-centre trial is to find out whether people with these more serious fractures do better if they have an operation within a few days of their injury. Surgeons and others, including people who have had these injuries, have agreed that the best way to find this out is to conduct a randomised trial. This means that the surgeon and patient agree to use whatever treatment (surgery or no surgery) is selected at random (or by chance). This is the best way to ensure that the two treatment groups are similar at the start. The results of this trial are likely to influence how these fractures are treated in this hospital and more generally in the future.

Why am I being asked to take part?

You have been invited to enter this trial because you have a serious break (fracture) at the top end of your upper arm bone. Your surgeon feels that you are fit enough for an operation but does not know if you would have just as good a result if you didn't have surgery.

Do I have to take part?

No. It is up to you to decide. You are free to withdraw at any time, without giving a reason. This would not affect the standard of care you receive.

What will happen to me if I take part?

After signing the consent form, you will be asked a few further questions. Then you will find out whether you have been allocated to surgery or not. In this trial there is an equal (or 50:50) chance of being allocated to surgery or no surgery.

You will then receive the allocated procedures delivered with the same standard of care that you would receive if you were not in the trial.

Thus, if you are allocated surgery, your surgeon will discuss with you the type of surgery he or she recommends for you. Your operation will be carried out by an orthopaedic surgeon experienced in treating shoulder fractures. Your surgeon will advise about the best care of your arm after your operation. Usually you will keep your arm in a sling for about three weeks. This will be followed by exercises.

If you are allocated to no surgery, your surgeon will again advise about the best non-operative treatment for you. Usually you will keep your arm in a sling for about three weeks, followed by exercises.

The only other thing that we ask of you is to complete a short questionnaire at three months and a longer one at six months, one year and two years after your injury. These questionnaires are very important and should not take longer than fifteen minutes to complete. You can ask a friend or relative to help you with these. A pre-paid envelope will be included. The information you give will be treated as confidential and only used for the trial.

You will need to tell us if you are happy for us to text you via your mobile number, if you have one. This would only be used to inform you of when to expect a questionnaire.

With your permission, we will inform your family doctor that you are taking part in the trial. Again with your permission, if we have difficulties contacting you we would like to ask your family doctor about whether it is appropriate to contact you and for your address.

We will also collect data from the hospital. Initially, this will be what treatment you received in hospital for your fracture and whether you have any complications. We will also ask the hospital to check their records after one and two years in case you had any later problems.

When we contact you at two years we will also ask if you would like to know the results of the trial when these become available.

What are the alternatives to taking part?

If you choose not to take part then your surgeon will discuss with you the options available for your treatment. It is possible that your surgeon may advise you to have surgery. Whatever you decide about taking part in the study it will not affect the standard of care you receive.

What are the possible disadvantages of taking part?

While the selection of treatment, surgery or no surgery, is left to chance you will receive the same standard of care as you would normally. The surgeon and other people providing your care are experienced in the treatments provided. There are no new treatments being tested in this study.

However, because all surgery involves extra risks, such as infection, it is very important that you consider whether you would be prepared to have surgery. You should also consider though that surgery may give you better shoulder function and may mean that you can be more independent. Sometimes too, surgery may be needed to correct problems later on. Such surgery is often more difficult to do and you may have a long wait.

If you receive an operation, you will usually be given a general anaesthetic. Although general anaesthesia is very safe with modern techniques, there are still very small risks involved. You may experience nausea, vomiting and / or dizziness. These are reduced with modern drugs. It is important that you tell the research team, medical staff and anaesthetist about any medical problems you have.

There is also a small chance of developing a wound infection. The exact risk is not known but most resolve with antibiotics. Occasionally, the metal implant may need to be removed. This will involve another operation. Another operation may also be needed if the fixation fails. For example, movement of the metal implant or parts of the implant may cause problems such as restricted movement and pain. Again, the exact risk of this happening is not known. You can be assured though that your surgeon will use devices with low complication rates and with which they are familiar.

It is important to realise that there are risks and problems for non-operative treatment too. If you don't have surgery, it may take longer for you to recover. It will mean that the fragments of bone will not be put back in place. This means that while the bone may heal, it will not be as it was before. This may give problems such as restricted movement and pain. You may start getting other shoulder problems too. Sometimes, these difficulties are bad enough to make you more dependent on the help of others. If this occurs, an operation may sometimes help. But because it is not a fresh fracture and other damage may have occurred, the surgery involved is usually more complex.

For both operative and non-operative treatment, some problems take a while to show themselves. This is why we ask people to tell us how they and their shoulder are at two years.

What are the possible benefits of taking part?

Because we do not know what treatment is best you might not benefit from taking part but by taking part you may help others in the future. We believe though this study has already increased awareness of injuries such as yours and recognition of the need for good quality care. If enough people take part in this study, the information we get should help ensure that people with these injuries have the best treatment in the future.

What will happen if I don't want to carry on with the study?

You are free to withdraw from the study at any time. You do not have to give a reason for this. Withdrawing will not affect your future care or rights in any way. It is up to you whether you tell us. However, if you do let us know of your decision then we will know not to contact you in future. It is also up to you to tell us whether you would prefer that we didn't contact the

hospital for any further outcome data and whether you would like your contact details to be deleted from the records. All the other data collected for you up to the time of your withdrawal from the study will be kept.

Expenses and payments

This trial is funded by the NHS and involves you in no extra treatment or tests. Thus no patient expenses have been allocated. However, in recognition of your help with this trial you will receive an unconditional payment of five pounds at one and two years after your injury when asked to complete questionnaires.

What will happen to data that are collected about me?

Your data will be held in a secure place in the co-ordinating centre in University of York. All data for the trial will be held for a minimum of 20 years. We will remove all names and other identifying information before the data are analysed and results of the trial presented to the medical community.

What happens if something goes wrong?

This trial only includes treatments that you would receive normally. The clinicians treating you will take every opportunity to reduce risk. If something were to go wrong, they will offer the best possible solution to resolve it. If you believe that you have been harmed in any way by taking part in this study, you have the right to pursue a complaint through the usual NHS procedures. Details about this will be available locally.

Who has reviewed this study?

Before any research goes ahead it has to be checked by a Research Ethics Committee. They make sure that the research is fair. This project was passed through the National Research Ethics Service as well as the local Ethics Committee for your hospital.

Who is organising and funding this research.


This trial is funded by the NHS R&D Health Technology Assessment Programme. The sponsor is the University of Teesside, Middlesbrough. Trial management is by the York Trials Unit (YTU), University of York. This trial has received endorsement by the British Elbow and Shoulder Society.

None of surgeons involved will receive payments for their involvement in the trial. The hospitals receive payments for entering a patient into the trial but these only cover the extra expenses incurred by the hospital for helping with this trial.

Who can I contact for more information?

If you have any queries or you wish to obtain further information about this study, please contact [designated local contact] on [telephone]. Alternatively you may contact Dr Stephen Brealey (PROFHER Trial Co-ordinator) on 01904 [redacted].

Thank you for reading this information sheet and for considering whether to take part in this study.

	<p>PROximal Fracture of Humerus: Evaluation by Randomisation (PROFHER) Trial <i>A multi-centre randomised controlled trial funded by NHS R&D Health Technology Assessment Programme (International Standardised Randomised Controlled Trial Number 50850043)</i></p>
---	---

Appendix 10 Payments for ProFHER trial-related activities

Hospitals

Activity per patient	Payment (£)
1. For patients meeting the inclusion criteria but subsequently judged non-eligible: completed study eligibility form	10
2. For eligible but non-consenting patient: completed study eligibility form and consent status form	30
3. For randomised patients: completed study eligibility form, consent status form, consent form, baseline form, treatment confirmation form and surgical form	500
4. Hospital data: provision of copies of radiographs and completed physiotherapy treatment log, physiotherapy treatment log – completion of treatment form and inpatient episode form	250
5. Hospital data: completed 1-year follow-up form and 2-year follow-up form	250

As noted in the trial protocol, the following activities were explicitly covered by these reimbursement payments:

- distribution of trial materials in appropriate locations
- setting up the trial processes
- screening of patients' notes for eligibility
- contacting patients or alerting appropriate clinicians for this activity
- discussing the trial with patients and providing documentation for consideration
- completion of trial eligibility and consent status forms
- obtaining consent
- randomisation of individual patients and completion of the baseline form
- completion of forms for treatment, costs and in-hospital outcomes and complications for randomised patients
- completion of forms for randomised patients at 1 and 2 years
- copying and carriage of baseline radiographs for independent review.

Patients participating in the trial

Patients did not receive payment for participation in the trial. An unconditional incentive payment of £5 was included in the questionnaires sent at 12 and 24 months' follow-up.

Trial Steering Committee, Data Monitoring and Ethics Committee and Trial Management Group meetings

Travel and subsistence expenses were provided only. When appropriate, standard refreshments and meals were provided.

Appendix 11 Recommended radiographic views for the ProFHER trial

Trauma series

What views?

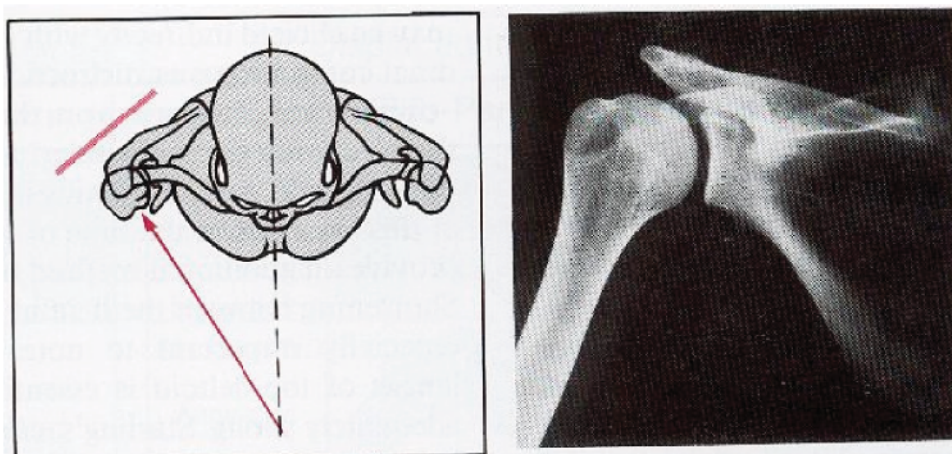
The following constitute the 'trauma series':

1. anteroposterior view (perpendicular to the scapular plane)
2. scapular Y-lateral view (parallel to the scapular plane)
3. axillary view (or modified axillary view).

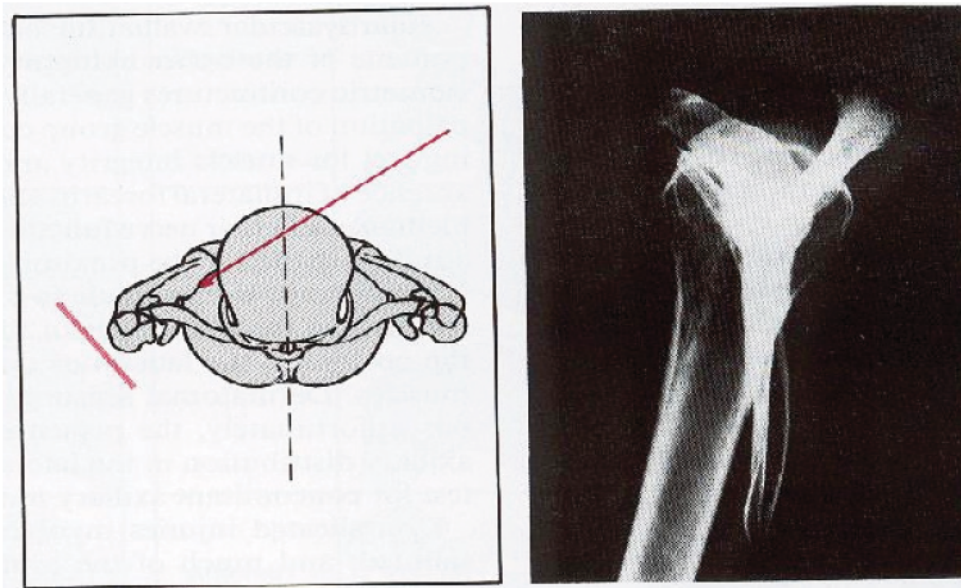
Why?

- recommended as good practice in the literature for evaluation of shoulder trauma
- a minimum of two radiographs in planes perpendicular to each other are required to safely plan management.

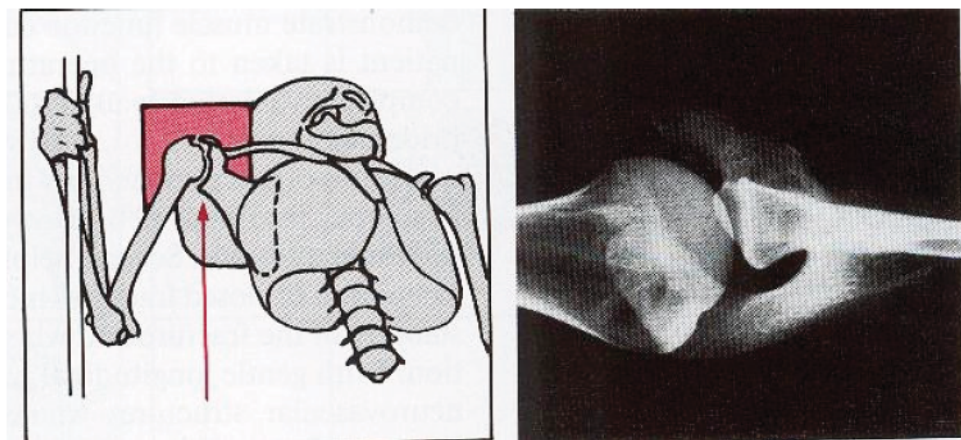
Anteroposterior view in the scapular plane



This figure was published in *Skeletal Trauma: Basic Science, Management and Reconstruction*, volume 2, Browner, BD, Jupiter JB, Levine AM, Trafton PG, editors, p. 1222, Copyright Elsevier, 1992. Reproduced with permission.

Scapular Y-lateral view

This figure was published in *Skeletal Trauma: Basic Science, Management and Reconstruction*, volume 2, Browner, BD, Jupiter JB, Levine AM, Trafton PG, editors, p. 1222, Copyright Elsevier, 1992. Reproduced with permission.

Axillary view

This figure was published in *Skeletal Trauma: Basic Science, Management and Reconstruction*, volume 2, Browner, BD, Jupiter JB, Levine AM, Trafton PG, editors, p. 1222, Copyright Elsevier 1992. Reproduced with permission.

If the axillary is not feasible/too uncomfortable for the patient, perform a modified axillary view (or Velpeau view), which can be carried out with the injured arm in a sling.

Appendix 12 The ProFHER trial sling immobilisation leaflet

Fractured Proximal Humerus Information for patients on initial self-care

You have been given this leaflet because the top end of your upper arm bone is broken. This is called a 'proximal humerus fracture'.

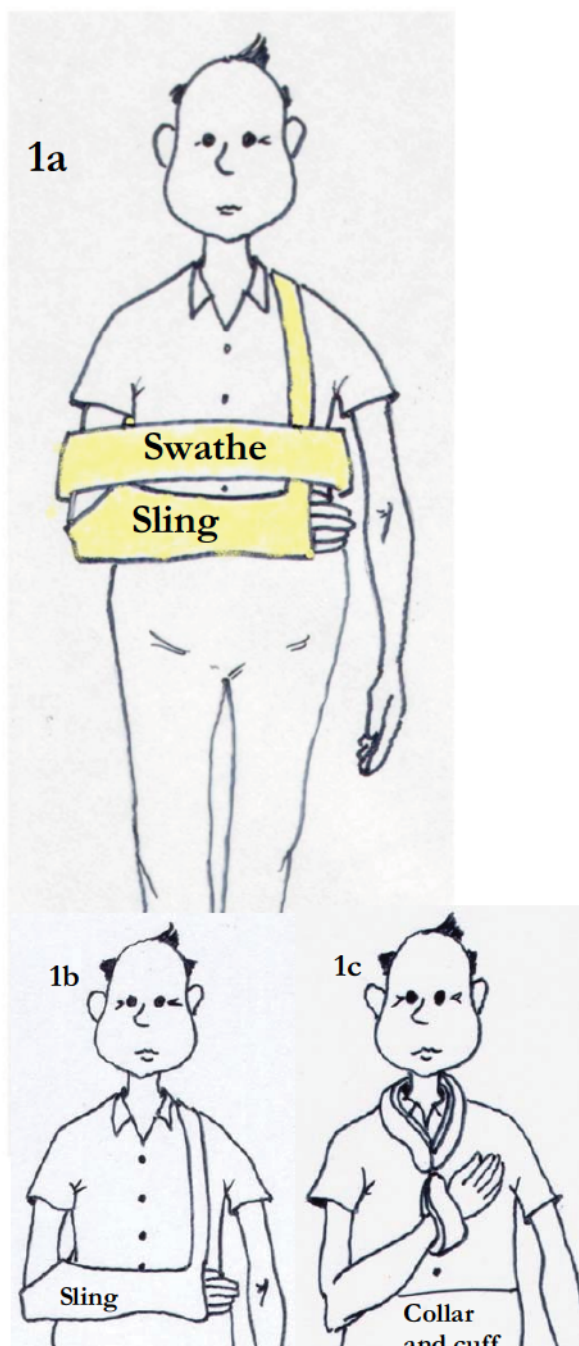
This leaflet is to remind you of the advice on self-care that you will receive from your hospital staff. Members of staff will be happy to explain any of the matters raised in this leaflet and you can also ask your family doctor (GP) for further advice when you have left hospital.

This leaflet covers the first few weeks after your injury when your arm is in a sling. People are usually advised to wear their sling for about three weeks. The sling will ease the pain and help the bone and soft tissues to heal, so it is important that you wear it both day and night. The sling should support the weight of your arm. In some hospitals, depending on the consultant's preference, the sling is secured by a 'swathe'. In others, a 'collar and cuff' is used instead of a sling. A well-positioned sling and swathe should look like diagram 1a. Diagram 1b shows a sling without a swathe, and diagram 1c shows a collar and cuff.

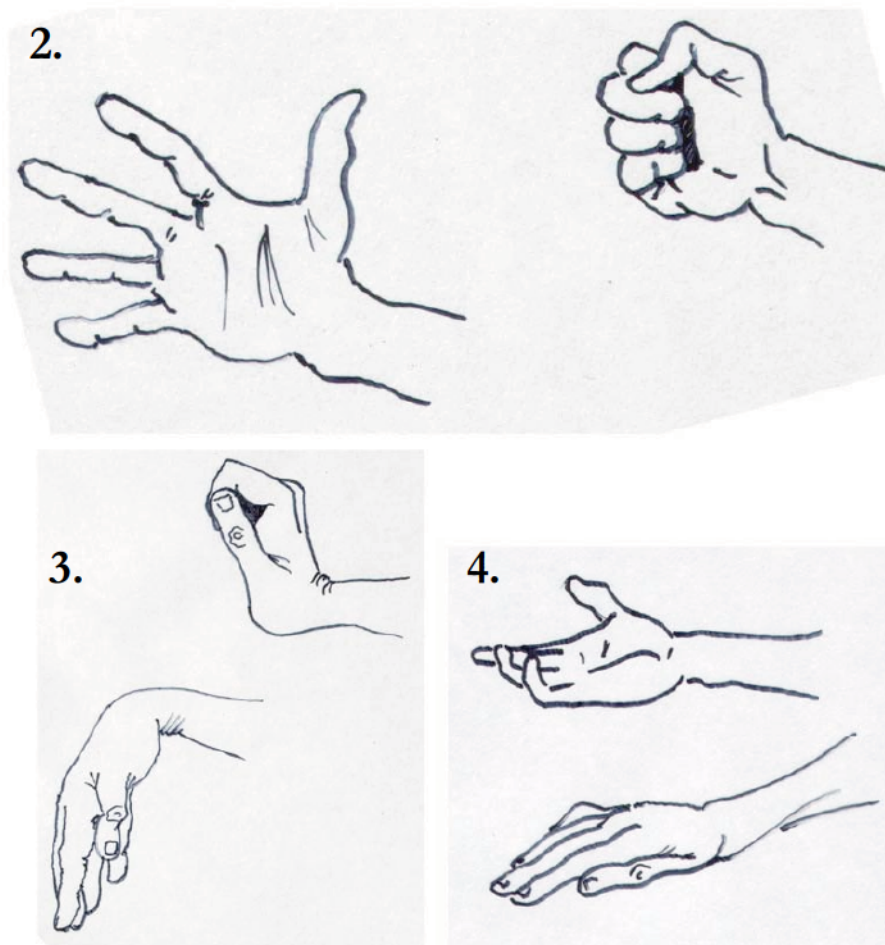
In the following sections we tell you things you should DO, including some tips on pain relief and on how you can make yourself more comfortable, things you should NOT DO, and things you MUST TELL YOUR HOSPITAL STAFF OR FAMILY DOCTOR (GP) ABOUT.

Do

- remove rings and watches from your arm as soon as possible after the injury, your hand and fingers might swell;
- make an effort to keep your wrist and fingers moving every couple of hours whilst you are wearing the sling, to prevent stiffness and to maintain your circulation. To do



this, open and close your fingers (diagram 2), and gently move your wrist ('backwards and forwards' (diagram 3), in a circular motion. Next, turn your palm up towards the ceiling then turn down towards the floor (diagram 4).

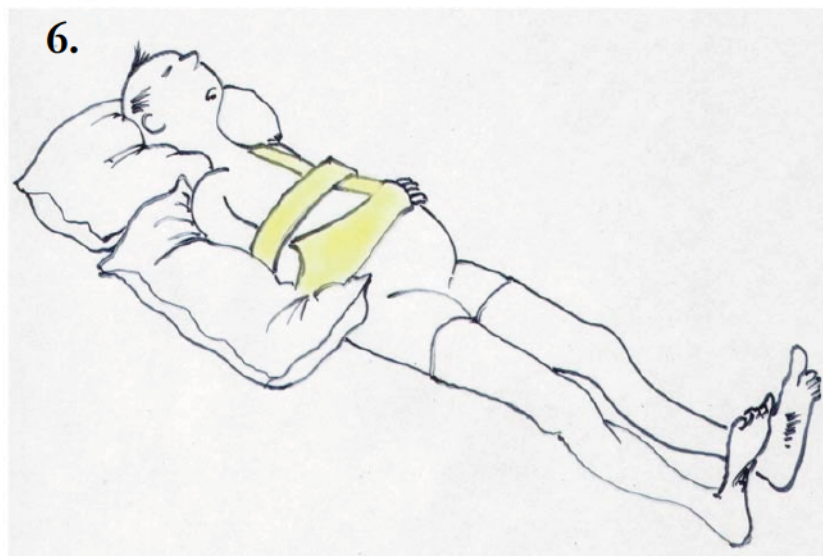


- check the sling regularly to ensure that it is not too tight;
- try to keep the sling as clean and dry as possible (you should ask for a replacement sling if yours becomes very grubby and uncomfortable);
- wear loose fitting clothing and front-fastening tops, as this will make it easier to dress and undress;
- learn to put your sling on and off by yourself, if possible;

- wash under your armpit (diagram 5) and in your elbow creases to help prevent these areas from becoming sore;



- support the arm with extra pillows behind your shoulder and also under your elbow when sleeping (diagram 6)



(You may find that you are more comfortable sleeping in a more upright position at first, gradually lying flatter as your pain settles.)

- discuss medication with your hospital doctor or family doctor (GP), as you may require regular pain relieving tablets until the pain subsides.
- be aware of the position of your shoulders when you are getting pain. If you feel tense, try taking a deep breath in, and then, when you breathe out, letting your shoulders relax. Being aware of your breathing may also help to reduce tension, as it is easy to hold your breath and forget to breathe correctly when you are in pain. Once you have got yourself supported with pillows in bed try closing your eyes, breathing slowly and deeply, and imagining pain and tension disappearing.

Do not

- use your arm for lifting or carry objects in your hand;
- drive (you will be advised when it will be safe for you to start driving again);
- operate machinery while wearing the sling;
- try to lie on your injured arm in bed at night (this is likely to be very uncomfortable);
- worry if you start to feel 'low', as this is something that many people experience with this type of injury. You could talk your feelings through with a friend or relative. If this does not help, you may wish to discuss any concerns with your GP.

Important:


Things you must tell your doctor about

If you suffer from any of the following please contact your doctor.

- Very bad pain that seems to be getting worse.
- Loss of sensation in the arm.
- Prolonged tingling or pins and needles in the arm.
- Fingers or hand turning cold and clammy or blue

Finally

Please do not hesitate to approach staff at your hospital for further explanations or if there is anything else you wish to know that has not been covered in this leaflet.

	<p>PROximal Fracture of Humerus: Evaluation by Randomisation (PROFHER) Trial <i>A multi-centre randomised controlled trial funded by NHS R&D Health Technology Assessment Programme (International Standardised Randomised Controlled Trial Number 50850043)</i></p>
---	---

Appendix 13 Questionnaire for physiotherapists on physiotherapy for proximal humeral fractures

Physiotherapy for patients with proximal humeral fractures: Questionnaire

The ProFHER trial, funded by the Department of Health's Health Technology Assessment Programme, is a UK-based, pragmatic, multi-centre randomised trial comparing surgical versus non-surgical treatment for displaced fractures of the proximal humerus. Important to the validity of this trial is the provision of comparable and good standard care including physiotherapy to both treatment groups. In our preparations to achieve this objective, we have developed a 'basic' physiotherapy protocol to guide treatment. The importance of this issue was emphasised by the Trial Steering Committee who recently (May 2008) recommended that feedback on the draft protocol and further insights into the current management of these injuries should be obtained from specialist shoulder physiotherapists and other experts in the field. This is the reason for this short questionnaire.

Please complete this form ONLY if you have ever treated people with shoulder injury.

Name: Specialty:

- Do you treat people with these fractures in your current practice? Yes / No
- If "No", have you ever treated patients with these fractures? Yes / No

Please answer the following questions in terms of the **expected trial population**: older people who are generally retired but still active and able to understand instructions for rehabilitation.

1. Physiotherapy protocol (please note this is a guide for the purposes of the trial and does not rule out the use of other modalities)

a. Is this acceptable as a basic treatment protocol for these fractures? Yes / No

If "No", please explain why not:

b. Would you be happy with following the approach given in the protocol? Yes / No

If "No", please explain why not:

2. Please list other interventions / modalities that you apply routinely for people with these fractures?

3. Please list any interventions available in an NHS setting that you would discourage for people with these fractures?

Please send your completed questionnaire by 23rd June to Dr Stephen Brealey at steve.brealey@york.ac.uk

Thank you very much for your help.



PROximal Fracture of Humerus: Evaluation by Randomisation (PROFHER) Trial
A multi-centre randomised controlled trial funded by NHS R&D Health Technology Assessment Programme
(International Standardised Randomised Controlled Trial Number 50850043)

Appendix 14 Physiotherapy protocol for the ProFHER trial

Upper limb specialist physiotherapists developed this protocol to promote standard care for patients in the ProFHER trial. It has received substantial endorsement from other specialist physiotherapists experienced in treating people with these fractures.

Patients participating in the ProFHER trial will have sustained an acute displaced fracture of the proximal humerus. The trial compares surgical versus non-surgical treatment for these fractures and it is important to the validity of the trial that comparable physiotherapy is provided to both groups of patients. These guidelines are not intended to replace local treatment protocols, but you must document any substantive difference from them in the accompanying physiotherapy treatment log.

We anticipate that home exercises will be the core of patients' therapy and that you might use a range of standard physiotherapy modalities/techniques to facilitate these, according to your clinical judgement. But please **do not** use any experimental or new treatment modalities/techniques. Also, with the exception of TENS, please **do not** use electrotherapy.

We have produced an information leaflet aimed at all patients with proximal humeral fractures. This includes advice on self-care and pain relief and describes the symptoms that should prompt patients to urgently seek medical advice. We envisage that this leaflet (or a locally developed leaflet covering the same material) will be handed out at the fracture clinic.

Treatment objective

The aims are to allow normal healing of bone and soft tissue and to aid the functional recovery of the injured arm.

Cautions

Please discuss any concerns you may have regarding possible complications with the treating consultant. Complications may include:

- avascular necrosis
- brachial plexus injury
- chest injury
- frozen shoulder
- malunion
- non-union and
- vascular injury.

This protocol is only a guideline and we accept that physiotherapists will use clinical reasoning and professional judgement when progressing rehabilitation for individual patients.

This protocol presents progression in phases. Their timing will depend on various factors including the patient's:

- age
- stage of healing
- pain tolerance
- expectation and
- general health and activity level.

With a stable fracture, phase 3 will probably start after removal of the sling at 3 weeks and phase 4 about 3 weeks later. It is anticipated that many patients will have returned to their former or attainable functional status by phase 6. If not, and a higher level of rehabilitation is required, completion may take several months.

All exercises must be performed within the limits of pain. Please advise patients to *do their home exercises regularly*, attempting at least five repetitions three times per day, and *not to force or stretch*.

Phase 1

- If the patient is seen early (i.e. on the ward) teach deep breathing exercises.
- Teach wrist and hand exercises to do whilst wearing the sling.
- Encourage maintenance of good posture for scapular stability.
- Gentle, slow, controlled shoulder shrugs and cervical range of motion exercises may help to alleviate muscle spasm.
- Teach the patient appropriate pain management techniques.
- Ensure the patient and/or his/her carer is taught sling application and axillary hygiene.

Phase 2

As for phase 1, except the sling may be removed for the patient to do elbow, wrist and hand exercises if the consultant feels it is safe for him or her to do so. Add:

- elbow flexion/extension and
- pronation/supination.

Phase 3

The consultant will recommend progression to this stage when radiographs show evidence of healing.

As for phase 2, except add:

- Gentle pendular exercises
- Active assisted flexion progressing from short to long lever and
- Active assisted external rotation (using a stick) in neutral with elbow flexed and supported on a pillow. Patients usually find this more comfortable to do in lying position.

Phase 4

Patients usually reach this stage in about 6 weeks. As for phase 3, except discontinue the hand and wrist exercises when appropriate and add:

- isometric internal and external rotation exercise in mid-range, within pain tolerance
- closed kinetic chain exercises
- light functional activities
- progression of active assisted exercises to active, gradually increasing range of motion, and
- *gentle* capsular stretches within pain tolerance if these are required.

Phase 5

As for phase 4, except add:

- progressive strengthening exercises and endurance activities for the rotator cuff muscles, ensuring optimum movement patterns and control through range of movement, and
- progression of functional activities as able.

Phase 6

Orthopaedic review is required to determine whether further progression is appropriate.

As for phase 5, except:

- progress strengthening exercises appropriate to patient's premorbid activity level and
- return to function/work/sport.

Discharge criteria

The patient is discharged:

- when independent shoulder function is achieved or
- by consensus between the therapist and patient if there has been no improvement over several sessions.

Appendix 15 Consent status form

Version 4.0 (23/09/2009)

ProFHER TRIAL: CONSENT STATUS FORM

Participant identification number:

The **designated person** should complete the first part of this form to indicate whether an **eligible patient** has consented to take part in the trial.

Name of designated person:

Date today: / /
Day Month Year

Has the patient agreed to consent?
(Please cross one box only)

Yes ☐No ☐

If the patient has **consented** then the designated person should complete the baseline form with the patient, perform the randomisation and then post all baseline materials to York Trials Unit.

If the patient has **not consented** then the **orthopaedic surgeon** should complete this form and give it to the designated person to post to York Trials Unit.

a. Which treatment do you as the **clinician** advise the patient to have?
(Please cross one box only)

Uncertain ☐Surgery ☐No Surgery ☐

b. Does the **patient** express any treatment preference?
(Please cross one box only)

No preference ☐Surgery ☐No Surgery ☐

c. What is the **agreed** treatment for this patient?
(Please cross one box only)

Surgery ☐No Surgery ☐

Thank you very much for completing this form



PROximal Fracture of Humerus: Evaluation by Randomisation (ProFHER) Trial
A multi-centre randomised controlled trial funded by NHS R&D Health Technology Assessment Programme
(International Standardised Randomised Controlled Trial Number 50850043)

4625359981

Appendix 16 Patient consent form

Version 1.0 (13/12/2007)

CONSENT FORM

Participant Identification Number:

--	--	--	--

Title of study: PROximal Fracture of Humerus: Evaluation by Randomisation
(ProFHER) Trial


Name of Person taking consent: _____

Please initial the boxes

1. I confirm that I have read and understand the information sheet version 3.0 dated 08/04/2008 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving reason, without my medical care or legal rights being affected. I also understand that if I withdraw, I can ask for all record of my contact details to be deleted but it will be impossible to withdraw any other data collected on me. ☐
3. I understand that sections of any of my medical notes and the data collected during the study may be looked at by responsible individuals where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐
4. I agree to my General Practitioner being informed that I am taking part in this study. ☐
5. Should there be problems contacting me, I agree to my General Practitioner being asked whether it is appropriate to contact me and for my address. ☐
6. I agree to this consent form and other data collected as part of this research study being kept at York Trials Unit, University of York. ☐
7. I understand that records relating to me will be kept confidential. No information will be released or printed that would identify me without my permission unless required by law. ☐
8. I agree to take part in the above study. ☐

Name of patient (please print)_____
Date_____
Signature_____
Name of person taking consent
(please print)_____
Date_____
Signature

When completed, 1 for patient; 1 (original) to be kept in medical notes; 1 for York Trials Unit

P R O F H E R		<p>PROximal Fracture of Humerus: Evaluation by Randomisation (PROFHER) Trial</p> <p><i>A multi-centre randomised controlled trial funded by NHS R&D Health Technology Assessment Programme (International Standardised Randomised Controlled Trial Number 50850043)</i></p>
--	---	--

Appendix 17 Baseline information form



Version 5.0 (22/08/2008)

THE UNIVERSITY of *York*

CONFIDENTIAL

PROximal Fracture of Humerus: Evaluation by Randomisation (ProFHER) Trial

**Baseline form for all patients with a radiologically confirmed
displaced fracture of the surgical neck**

Participant ID Number:

Date when questionnaire
completed:

day

month

year



PROximal Fracture of Humerus: Evaluation by Randomisation (ProFHER) Trial
A multi-centre randomised controlled trial funded by NHS R&D Health Technology Assessment Programme
(International Standardised Randomised Controlled Trial Number 50850043)

9457133800

BASELINE FORM FOR CONSENTING PATIENTS

This form is for the designated person to complete with eligible patients who consent to take part in the trial.

Section 1:

This section asks about the patient's health in general. Please indicate which statement best describes the patient's health state today.

Please cross one box in each group.

1. Mobility

I have no problems in walking about

☐

I have some problems in walking about

☐

I am confined to bed

☐

2. Self-care

I have no problems with self-care

☐

I have some problems washing or dressing myself

☐

I am unable to wash or dress myself

☐

3. Usual activities (e.g. work, study, housework, family or leisure activities)

I have no problems with performing my usual activities

☐

I have some problems with performing my usual activities

☐

I am unable to perform my usual activities

☐

4. Pain or discomfort

I have no pain or discomfort

☐

I have moderate pain or discomfort

☐

I have extreme pain or discomfort

☐

5. Anxiety or depression

I am not anxious or depressed

☐

I am moderately anxious or depressed

☐

I am extremely anxious or depressed

☐

7850133805

Section 2: IT WOULD ALSO HELP TO KNOW MORE ABOUT YOU

1. Are you?
(Please cross one box)
- | | |
|---------|--------------------------|
| White | <input type="checkbox"/> |
| Black | <input type="checkbox"/> |
| Asian | <input type="checkbox"/> |
| Chinese | <input type="checkbox"/> |
| Other | <input type="checkbox"/> |

If 'Other', please specify here:

2. Which of these best describes your highest qualification?
(Please cross one box)
- | | |
|-------------------------------|--------------------------|
| No formal qualifications | <input type="checkbox"/> |
| Some qualifications/no degree | <input type="checkbox"/> |
| Degree or higher | <input type="checkbox"/> |

3. Are you?
(Please cross the box that describes you best)
- | | |
|-------------------------------|--------------------------|
| Employed part-time | <input type="checkbox"/> |
| Employed full-time | <input type="checkbox"/> |
| Self-employed | <input type="checkbox"/> |
| Student | <input type="checkbox"/> |
| Retired | <input type="checkbox"/> |
| Not employed but seeking work | <input type="checkbox"/> |
| Other | <input type="checkbox"/> |

If 'Other', please specify here:

7134133800

4. Have you had any other fractures in the past ten years? Yes ☐
 (Please cross one box)

No ☐

If you have placed a cross in the 'Yes' box, please record:

What type of fracture(s)?

Was it/were any treated surgically?
 (Please cross one box)

Yes ☐

No ☐

5. Do you smoke?
 (Please cross one box)

Yes ☐

No ☐

If you have placed a cross in the 'Yes' box, please be advised that smoking has been shown to delay bone healing and recovery and thus it would be a good time to stop smoking if at all possible.

6. How did you injure your shoulder?
 (Please cross one box)

Fall/trip from standing height or less ☐

Fall downstairs/steps or from a height ☐

Other (please describe in box below) ☐

7. Have you injured your dominant arm?
 (Please cross one box)

Yes ☐

No ☐

8. Are you diabetic?
 (Please cross one box)

Yes ☐

No ☐

9. Are you currently taking steroids?
 (Please cross one box)

Yes ☐

No ☐

10. In consenting for this trial, you have accepted that whether you get surgery or no surgery is left to chance. However, we would like to know if you had a preference before you agreed to this?
 (Please cross one box)

No preference ☐

Surgery ☐

Not Surgery ☐

4521133807

11. What is the name of your GP? *(Please record below)*

12. What is the name of your GP practice (include postcode if available)?
(Please record below)

13. Are you happy to receive a text about when to expect a questionnaire?
(Please cross one box)

Yes ☐

No ☐

The rest of the form can be completed by the designated person without the patient.

14. Please record below the patient's contact details:

Patient's Name	<input type="text"/>
Address Line 1	<input type="text"/>
Address Line 2	<input type="text"/>
Address Line 3	<input type="text"/>
Postcode	<input type="text"/>
Telephone (Day)	<input type="text"/>
Telephone (Evening)	<input type="text"/>
Mobile	<input type="text"/>
Email:	<input type="text"/>

15. Please randomise the patient by contacting York Trials Unit on freephone 0800 XXXXXXXXXX or the following website www.yorkrand.com

16. Please cross the box for which treatment the patient has been randomised to receive:

Surgery ☐

Not Surgery ☐

Please inform the patient about the treatment they will receive.

Please post this baseline form to York Trials Unit with the signed consent form, consent status form and study eligibility form using the freepost envelope provided.

Thank you very much for completing this form.

3369133806

Appendix 18 Three-month patient questionnaire



Version 3.0 (01/07/2008)

THE UNIVERSITY of York

CONFIDENTIAL

PROximal Fracture of Humerus: Evaluation by Randomisation (ProFHER) Trial

**Three month questionnaire to find out how you are and
your health care needs after your shoulder fracture**

Office use only

Participant ID Number:

--	--	--	--

Date when questionnaire
sent:

		/			/				
day			month			year			



PROximal Fracture of Humerus: Evaluation by Randomisation (ProFHER) Trial
A multi-centre randomised controlled trial funded by NHS R&D Health Technology Assessment Programme
(International Standardised Randomised Controlled Trial Number 50850043)

8688648716

**PLEASE READ ALL THE INSTRUCTIONS BEFORE COMPLETING
THE QUESTIONNAIRE**

Thank you for agreeing to take part in this study.

The responses you give in this questionnaire will help us to find out if the care you receive is helpful for your shoulder fracture.

Please answer **all** the questions.

If you find it difficult to answer a question, please give the best answer you can.

Please follow the instructions for each section carefully.

Please use a black or blue pen for all of the questions.

Please do not use a pencil or any other coloured pen.

Please read all the instructions for each section.

The responses you give will be treated in the utmost confidence. The study team will not tell anyone else what you have written.

Please enter today's date:

<input type="text"/>	<input type="text"/>	/	<input type="text"/>	<input type="text"/>	/	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
day			month			year			

Section 1:

This section asks about your health in general. By placing a cross in one box in each group below, please indicate which statement best describes your health state today.

Please cross one box in each group.

1. Mobility

I have no problems in walking about

☐

I have some problems in walking about

☐

I am confined to bed

☐

2. Self-care

I have no problems with self-care

☐

I have some problems washing or dressing myself

☐

I am unable to wash or dress myself

☐

3. Usual activities (e.g. work, study, housework, family or leisure activities)

I have no problems with performing my usual activities

☐

I have some problems with performing my usual activities

☐

I am unable to perform my usual activities

☐

4. Pain or discomfort

I have no pain or discomfort

☐

I have moderate pain or discomfort

☐

I have extreme pain or discomfort

☐

5. Anxiety or depression

I am not anxious or depressed

☐

I am moderately anxious or depressed

☐

I am extremely anxious or depressed

☐

4986648713

This **final section** is about the health care you have received, over the **past three months** FOR ANY REASON. If the health care you received was related to your shoulder injury, record this in the '**about your shoulder**' column. If the health care was for any other reason, enter this in the '**other reason**' column. Please answer every question, even if the answer is "0".

Please fill in both boxes, for example: if seen three times.

0	3
---	---

Care from the NHS

In this part we would like you to tell us about the care you have received from the NHS.

1. Over the **past three months**, how many times have you:

	About your shoulder	Other reason				
a. Seen your GP? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
b. Seen a physiotherapist? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
c. Seen a nurse at your GP practice? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
d. Seen a district/community nurse? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
e. Seen an occupational therapist? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
f. Visited hospital for an out-patient appointment? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
g. Visited hospital as a day case (not overnight)? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
h. Stayed in hospital as an in-patient? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
i. If you have stayed in hospital, could you please record how many nights you were there over all visits.	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		

4584648717

Private treatments

In this part we would like you to tell us about **any additional medical treatments you have received, which you have paid for.**

2. Over the **past three months**, how many times have you:

	About your shoulder	Other reason
a. Seen a non-NHS physiotherapist? (Please record the number of times in the boxes)	<div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <i>If none enter '0'</i>	<div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <i>If none enter '0'</i>
b. Seen an osteopath? (Please record the number of times in the boxes)	<div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <i>If none enter '0'</i>	<div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <i>If none enter '0'</i>
c. Seen a private hospital doctor? (Please record the number of times in the boxes)	<div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <i>If none enter '0'</i>	<div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <i>If none enter '0'</i>
d. Seen a chiropractitioner? (Please record the number of times in the boxes)	<div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <i>If none enter '0'</i>	<div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <i>If none enter '0'</i>
e. Stayed in a private hospital? (Please record the number of times in the boxes)	<div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <i>If none enter '0'</i>	<div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <div style="display: inline-block; border: 1px solid black; width: 30px; height: 20px; margin-right: 5px;"></div> <i>If none enter '0'</i>
f. If you would like to tell us about any other medical treatments you have paid for, please record details below.		

Usual activities

3. Are you in any form of paid employment? Yes ☐ No ☐

If you have answered 'Yes' to being in any form of paid employment, how many working days over the **past three months** have you missed because of your shoulder? days

4. For how many days over the **past three months** have you been unable to perform your normal UNPAID activities (e.g. household chores, shopping, helping others) because of **your shoulder**? days

7936648711

Please post the questionnaire back to us using the freepost envelope provided.

We will ask you to complete a further questionnaire in three months time. Thus we would be most grateful if you could inform us of any change in your contact details.

If you have been admitted to a different hospital than the one you initially attended for your shoulder fracture, could you please tell us the name of the hospital.

Please let us know of anything that you think we have not asked that is badly affecting your everyday activities because of your shoulder.

If you have any other comments about your shoulder problem, this study, or this questionnaire, please write them here.

THANK YOU VERY MUCH FOR COMPLETING THIS QUESTIONNAIRE.

0717648711

Appendix 19 Two-year patient questionnaire



Version 1.0 (28/07/2010)

THE UNIVERSITY *of York*

CONFIDENTIAL

PROximal Fracture of Humerus: Evaluation by Randomisation (ProFHER) Trial

Two year questionnaire to find out how you are and
your health care needs after your shoulder fracture

Office use only

Participant ID Number:

Date when questionnaire
sent:

day

month

year



PROximal Fracture of Humerus: Evaluation by Randomisation (ProFHER) Trial
A multi-centre randomised controlled trial funded by NHS R&D Health Technology Assessment Programme
(International Standardised Randomised Controlled Trial Number 50850043)

7695553965

**PLEASE READ ALL THE INSTRUCTIONS BEFORE COMPLETING
THE QUESTIONNAIRE**

Thank you for agreeing to take part in this study.

The responses you give in this questionnaire will help us to find out if the care you received was helpful for your shoulder fracture.

Please answer **all** the questions. Although some questions may appear similar, it is still important that you answer every one.

If you find it difficult to answer a question, please give the best answer you can.

Please follow the instructions for each section carefully.

For sections where you are asked to put a cross in the box, please do this as follows. For example, if your answer to the following question is **No**, you should place a cross firmly in the box next to **No**.

Do you drive a car? Yes ☐

 No ☒

Please use a black or blue pen for all of the questions.

Please do not use a pencil or any other coloured pen.

Please read all the instructions for each section.

The responses you give will be treated in the utmost confidence. The study team will not tell anyone else what you have written.

Please enter today's date:

<input type="text"/>	<input type="text"/>	/	<input type="text"/>	<input type="text"/>	/	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
day			month			year			

YOUR SHOULDER PROBLEM

The following questions ask about the problems that the shoulder you had injured has caused you **during the past 4 weeks**. Please answer every question with a cross. If you are unsure about how to answer a question, please give the best answer you can.

Which shoulder did you injure?
(Please cross one box only)

Left ☐ Right ☐

1. During the **past 4 weeks**, how would you describe the worst pain you had from your injured shoulder?

(please cross one box only)

None	Mild	Moderate	Severe	Unbearable
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2. During the **past 4 weeks**, have you had any trouble dressing yourself because of your injured shoulder?

(please cross one box only)

No trouble at all	Little trouble	Moderate trouble	Extreme difficulty	Impossible to do
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3. During the **past 4 weeks**, have you had any trouble getting in and out of a car or using public transport because of your injured shoulder? (whichever you tend to use)

(please cross one box only)

No trouble at all	Very little trouble	Moderate trouble	Extreme difficulty	Impossible to do
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4. During the **past 4 weeks**, have you been able to use a knife and fork - at the same time?

(please cross one box only)

Yes, easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9063553966

5. During the **past 4 weeks**, could you do the household shopping on your own?

(Please cross one box only)

Yes, easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6. During the **past 4 weeks**, could you carry a tray containing a plate of food across a room?

(Please cross one box only)

Yes, easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7. During the **past 4 weeks**, could you brush/comb your hair with the affected arm?

(Please cross one box only)

Yes, easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8. During the **past 4 weeks**, how would you describe the pain you usually had from your injured shoulder?

(Please cross one box only)

None	Very mild	Mild	Moderate	Severe
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9. During the **past 4 weeks**, could you hang your clothes up in a wardrobe, using the affected arm?

(Please cross one box only)

Yes, easily	With little difficulty	With moderate difficulty	With great difficulty	No, impossible
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10. During the **past 4 weeks**, have you been able to wash and dry yourself under both arms?

(Please cross one box only)

Yes, easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

11. During the **past 4 weeks**, how much has pain from your injured shoulder interfered with your usual work (including housework)?

(Please cross one box only)

Not at all	A little bit	Moderately	Greatly	Totally
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

12. During the **past 4 weeks**, have you been troubled by pain from your injured shoulder in bed at night?

(Please cross one box only)

No nights	Only 1 or 2 nights	Some nights	Most nights	Every night
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4253553962

YOUR GENERAL HEALTH

The following questions ask for your views about your health. This information will help us keep track of how you feel and how well you are able to do your usual activities.

Answer each question by marking a cross in the box that best describes your answer.

1. In general, would you say your health is:

(Please cross one box only)

Excellent

☐

Very Good

☐

Good

☐

Fair

☐

Poor

☐

2. During a typical day does **your health** limit you in **moderate activities**, such as moving a table, pushing a vacuum cleaner, bowling or playing golf? If so, how much?

(Please cross one box only)

Yes, limited a lot

☐

Yes, limited a little

☐

No, not limited at all

☐

3. During a typical day does **your health** limit you in climbing **several** flights of stairs? If so, how much?

(Please cross one box only)

Yes, limited a lot

☐

Yes, limited a little

☐

No, not limited at all

☐

4. During the **past 4 weeks**, how much of the time have you accomplished less than you would like in regular daily activities **as a result of your physical health**?

(Please cross one box only)

All of the time

☐

Most of the time

☐

Some of the time

☐

A little of the time

☐

None of the time

☐

4647553969

5. During the **past 4 weeks**, how much of the time have you been limited in performing any kind of work or other regular daily activities **as a result of your physical health**?

(Please cross one box only)

All of the time	Most of the time	Some of the time	A little of the time	None of the time
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6. During the **past 4 weeks**, how much of the time have you accomplished less than you would have liked in your work or any other regular daily activities **as a result of any emotional problems** (such as feeling depressed or anxious)?

(Please cross one box only)

All of the time	Most of the time	Some of the time	A little of the time	None of the time
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7. During the **past 4 weeks**, how much of the time have you done work or other activities less carefully than usual **as a result of any emotional problems** (such as feeling depressed or anxious)?

(Please cross one box only)

All of the time	Most of the time	Some of the time	A little of the time	None of the time
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8. During the **past 4 weeks**, how much did **pain** interfere with your normal work (both outside the home and housework)?

(Please cross one box only)

Not at all	A little bit	Moderately	Quite a bit	Extremely
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5831553965

9. This question is about how you feel and how things have been with you during the **past 4 weeks**. Please give the one answer that comes closest to the way you have been feeling. How much of the time during the **past 4 weeks** have you felt calm and peaceful?

(Please cross one box only)

All of the time	Most of the time	Some of the time	A little of the time	None of the time
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10. This question is about how you feel and how things have been with you during the **past 4 weeks**. Please give the one answer that comes closest to the way you have been feeling. How much of the time during the **past 4 weeks** did you have a lot of energy ?

(Please cross one box only)

All of the time	Most of the time	Some of the time	A little of the time	None of the time
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

11. This question is about how you feel and how things have been with you during the **past 4 weeks**. Please give the one answer that comes closest to the way you have been feeling. How much of the time during the **past 4 weeks** have you felt downhearted and depressed?

(Please cross one box only)

All of the time	Most of the time	Some of the time	A little of the time	None of the time
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

12. During the **past 4 weeks** how much of the time has your **physical health** or **emotional problems** interfered with your social activities (like visiting friends, relatives etc.)?

(Please cross one box only)

All of the time	Most of the time	Some of the time	A little of the time	None of the time
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2385553966

This section also asks about your health in general. By placing a cross in one box in each group below, please indicate which statement best describes your health state today.

Please cross one box in each group.

1. Mobility

I have no problems in walking about

☐

I have some problems in walking about

☐

I am confined to bed

☐

2. Self-care

I have no problems with self-care

☐

I have some problems washing or dressing myself

☐

I am unable to wash or dress myself

☐

3. Usual activities (e.g. work, study, housework, family or leisure activities)

I have no problems with performing my usual activities

☐

I have some problems with performing my usual activities

☐

I am unable to perform my usual activities

☐

4. Pain or discomfort

I have no pain or discomfort

☐

I have moderate pain or discomfort

☐

I have extreme pain or discomfort

☐

5. Anxiety or depression

I am not anxious or depressed

☐

I am moderately anxious or depressed

☐

I am extremely anxious or depressed

☐

This **final section** is about the health care you have received, over the **past twelve months** FOR ANY REASON. If the health care you received was related to your shoulder injury, record this in the '**about your shoulder**' column. If the health care was for any other reason, enter this in the '**other reason**' column. Please answer every question, even if the answer is "0".

Please fill in both boxes: for example, if seen three times.

0	3
---	---

Care from the NHS

In this part we would like you to tell us about the care you have received from the **NHS**.

1. Over the **past twelve months**, how many times have you:

	About your shoulder	Other reason				
a. Seen your GP? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
b. Seen a physiotherapist? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
c. Seen a nurse at your GP practice? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
d. Seen a district/community nurse? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
e. Seen an occupational therapist? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
f. Visited hospital for an out-patient appointment? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
g. Visited hospital as a day case (not overnight)? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
h. Stayed in hospital as an in-patient? (Please record the number of times in the boxes)	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		
i. If you have stayed in hospital, could you please record how many nights you were there over all visits.	<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'			<table border="1"><tr><td></td><td></td></tr></table> If none enter '0'		

0774553966

Private treatments

In this part we would like you to tell us about **any additional medical treatments you have received, which you have paid for.**

2. Over the **past twelve months**, how many times have you:

	About your shoulder	Other reason
a. Seen a non-NHS physiotherapist? (Please record the number of times in the boxes)	<div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black; margin-right: 5px;"></div> <div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black;"></div> <i>If none enter '0'</i>	<div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black; margin-right: 5px;"></div> <div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black;"></div> <i>If none enter '0'</i>
b. Seen an osteopath? (Please record the number of times in the boxes)	<div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black; margin-right: 5px;"></div> <div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black;"></div> <i>If none enter '0'</i>	<div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black; margin-right: 5px;"></div> <div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black;"></div> <i>If none enter '0'</i>
c. Seen a private hospital doctor? (Please record the number of times in the boxes)	<div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black; margin-right: 5px;"></div> <div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black;"></div> <i>If none enter '0'</i>	<div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black; margin-right: 5px;"></div> <div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black;"></div> <i>If none enter '0'</i>
d. Seen a chiropractitioner? (Please record the number of times in the boxes)	<div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black; margin-right: 5px;"></div> <div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black;"></div> <i>If none enter '0'</i>	<div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black; margin-right: 5px;"></div> <div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black;"></div> <i>If none enter '0'</i>
e. Stayed in a private hospital? (Please record the number of times in the boxes)	<div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black; margin-right: 5px;"></div> <div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black;"></div> <i>If none enter '0'</i>	<div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black; margin-right: 5px;"></div> <div style="display: inline-block; width: 40px; height: 20px; border: 1px solid black;"></div> <i>If none enter '0'</i>

- f. If you would like to tell us about any other medical treatments you have paid for, please record details below.

If you have been admitted to a different hospital than the one you initially attended for your shoulder fracture, could you please tell us the name of the hospital.

5413553961

Other questions

1. **Compared with one year ago** how is your shoulder now?
(Please cross one box only)

**Much better
now than one
year ago**

☐

**Slightly better
now than one
year ago**

☐

**About the
same now as
one year ago**

☐

**Slightly worse
now than one
year ago**

☐

**Much worse
now than one
year ago**

☐

2. Based upon your experiences of the treatment that you received as part of this trial, if you injured your shoulder today to the same extent as you did two years ago which treatment would you prefer?
(Please cross one box only)

No preference ☐

Surgery ☐

Not surgery ☐

3. Would you like us to inform you about the results of this study?
(Please cross one box only)

Yes ☐

No ☐

Please let us know of anything that you think we have not asked that is badly affecting your everyday activities because of your shoulder.

If you have any other comments about your shoulder problem, this study, or this questionnaire, please write them here.

THANK YOU VERY MUCH FOR COMPLETING THIS QUESTIONNAIRE.

Please post the questionnaire back to us using the freepost envelope provided.

0773553965

Appendix 20 Treatment confirmation form

ProFHER TRIAL: One Month Treatment Confirmation Form

Participant identification number:

This form is for the **designated person** to complete at one month after trial entry.

1. Which treatment did the patient have?

(Please cross one box only)

☐

Surgery

☐

Not Surgery

2. If 'not surgery', please record the non-surgical treatment prescribed for this patient

(e.g. sling immobilisation):

3. If the patient had surgery, please record date of surgery:

Day

Month

Year

Check for completion of surgical form and confirm this (please cross one box only):

Completed surgical form:

☐

available

☐

not available*

4. Whether the patient was treated surgically or not, please record the following:

a. Was the patient admitted as an inpatient?

(please cross one box only)

☐

Yes

☐

No

If **Yes** and patient has left hospital, please check for completion of inpatient episode form or complete this (please cross one box only)

Completed inpatient episode form:

☐

available

☐

not available*

b. Please record date that patient started or is scheduled to start physiotherapy:

Day

Month

Year

If **no physiotherapy planned**, please give reason:

5. X-rays (at baseline)

Copies of the X-rays used to determine patient eligibility for the trial should be or have been sent to the trial office. Please take agreed steps to organise this.

Please record if copies have been sent / are being sent to the trial office:

(please cross one box only)

☐

Yes

☐

No

If X-rays **not** available, please contact trial office to discuss.

Please record in the box below any other imaging investigations (e.g. CT scan) used by the orthopaedic surgeon to decide on the patient's eligibility for the trial.

* If a form is **not** available, please contact trial office to discuss.

Thank you for completing this form.

8900062331

Appendix 21 Surgical form

ProFHER TRIAL: SURGICAL FORM

Participant identification number:

Date of surgery: / /

Day

Month

Year

Name of operating surgeon:

Operation - Times

hr min

hr min

hr min

Into anaesthetic room

Into theatre

Out of theatre

Type of Anaesthesia: (Please cross all boxes that apply)

General Anaesthetic ☐

Local Anaesthetic ☐

Regional Anaesthesia (RA) ☐

Other Anaesthesia ☐

If 'Other Anaesthesia' please record in the box

Delivery of Regional Anaesthesia: (Please cross one box only)

a) Indwelling catheter ☐

b) Injection ☐

Type of Surgery: (Please describe type of implant and stick implant label on back of the form)

Nail

Plate & Screws

Hemiarthroplasty

Other

Was a loan set/kit used for the surgery?

(Please cross one box only)

Yes ☐

No ☐

Use of antibiotics:

(Please cross one box only)

Yes ☐

No ☐

Staff in theatre: (Please record number of staff)

Surgeon grade:

Consultant

Registrar

Other

If Other grade, please record:

Anaesthetist grade:

Consultant

Registrar

Other

If Other grade, please record:

Other staff:
(e.g. Nursing)

Grade

Number of staff

Grade

Number of staff

Radiology: (Please cross all boxes that apply)

Image intensifier

Films

Radiographer grade (please record below)

☐
☐

Unexpected procedures: (Please cross one box only)

Yes ☐

No ☐

If yes, please describe and record reason:

Thank you for completing the form. The designated person should now return the form to York Trials Unit in the freepost envelope provided.

5706096485

Appendix 22 Inpatient episode form

ProFHER TRIAL: END OF ORTHOPAEDIC INPATIENT EPISODE FORM

Participant identification number:

This form is for the **designated person** to complete when an eligible patient comes to the end of their inpatient episode in the Orthopaedic Department.

Date of admission:

 / /

Day Month Year

Date of discharge:

 / /

Day Month Year

Date of death (if applicable)

 / /

Day Month Year

If the patient has died, please complete an **adverse event form** and return to York Trials Unit.

Discharged to: (please cross one box)

Home ☐ Another Hospital ☐ Nursing Home ☐ Other e.g. ward ☐

If Other, please give details:

Surgical and other shoulder fracture complications prior to discharge:

Please cross in box(es) if recorded

Surgical site infection requiring treatment with antibiotics

☐

Identified organism:

Haematoma formation at surgical site

☐

Describe:

Nerve injury

☐

Describe:

Implant related complication

☐

Describe:

Dislocation/instability

☐

Describe:

Other

☐

Describe:

Please continue to complete the rest of this form overleaf.

3356038659

Medical complications prior to discharge:*Please cross in box(es) if recorded*

Confirmed myocardial infarction

☐

Confirmed stroke

☐

Chest infection

☐

Other hospital acquired infection

☐

If Other hospital acquired infection, please describe:

Deep vein thrombosis requiring treatment

☐

Pulmonary embolism requiring treatment

☐

Admission to intensive care unit

☐

Number of days admitted to intensive care unit:

Admission to high dependency unit

☐

Number of days admitted to high dependency unit:

Other serious event

☐

If 'Other serious event', please describe:

Other treatments (after patient's initial treatment):

Secondary surgery to shoulder

Yes

☐

No

☐

Date:

Day

Month

Year

If 'Yes', please describe surgery in the box below:

Increased/new therapy for shoulder related complications (e.g. infection)

Yes

☐

No

☐

Date

started:

Day

Month

Year

If 'Yes', please describe treatment in the box below:

Treatment for serious newly diagnosed medical complication

Yes

☐

No

☐

Date

started:

Day

Month

Year

If 'Yes', please describe treatment in the box below:

Thank you for completing this form.

7338038653

Appendix 23 Physiotherapy treatment log

ProFHER TRIAL: PHYSIOTHERAPY TREATMENT LOG

Participant identification number: Sheet number:

Please complete this form as soon as possible after each treatment session. Thank you.

	Day	Month	Year	Day	Month	Year
Date:	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Location:	<input type="text"/>			<input type="text"/>		
Name of person completing form:	<input type="text"/>			<input type="text"/>		
Staff grade:	<input type="text"/>			<input type="text"/>		
Session number:	<input type="text"/>			<input type="text"/>		
Allocated time (mins):	<input type="text"/>			<input type="text"/>		
Phase*:	<input type="text"/>			<input type="text"/>		

*Please see ProFHER physiotherapy protocol. Note: phase 3 starts after sling removal.

Please record with a cross all that apply to the session:

Advice and/or education	<input type="checkbox"/>	<input type="checkbox"/>
Exercise	<input type="checkbox"/>	<input type="checkbox"/>
TENS	<input type="checkbox"/>	<input type="checkbox"/>
Soft tissue techniques	<input type="checkbox"/>	<input type="checkbox"/>
Joint mobilisations	<input type="checkbox"/>	<input type="checkbox"/>
Stretching techniques	<input type="checkbox"/>	<input type="checkbox"/>
Relaxation techniques	<input type="checkbox"/>	<input type="checkbox"/>
Hydrotherapy	<input type="checkbox"/>	<input type="checkbox"/>

Other:
(Please specify) Has the patient done their home exercises? ☐ Yes ☐ No (any comments*)☐ Yes ☐ No (any comments*)

Please record reasons below for any substantial difference from the ProFHER physiotherapy protocol.

Reasons
(or comments*)If referral to other speciality is required, please record speciality and reason

Please continue to complete the rest of this form overleaf for the next two sessions.

Prof_Physio_Treat_Logv2_Feb09

5166011303

Please complete this form as soon as possible after each treatment session. Thank you.

	Day	Month	Year	Day	Month	Year
Date:	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Location:	<input type="text"/>			<input type="text"/>		
Name of person completing form:	<input type="text"/>			<input type="text"/>		
Staff grade:	<input type="text"/>			<input type="text"/>		
Session number:	<input type="text"/>			<input type="text"/>		
Allocated time (mins):	<input type="text"/>			<input type="text"/>		
Phase*:	<input type="text"/>			<input type="text"/>		

*Please see ProFHER physiotherapy protocol. Note: phase 3 starts after sling removal.

Please record with a cross all that apply to the session:

Advice and/or education	<input type="checkbox"/>	<input type="checkbox"/>
Exercise	<input type="checkbox"/>	<input type="checkbox"/>
TENS	<input type="checkbox"/>	<input type="checkbox"/>
Soft tissue techniques	<input type="checkbox"/>	<input type="checkbox"/>
Joint mobilisations	<input type="checkbox"/>	<input type="checkbox"/>
Stretching techniques	<input type="checkbox"/>	<input type="checkbox"/>
Relaxation techniques	<input type="checkbox"/>	<input type="checkbox"/>
Hydrotherapy	<input type="checkbox"/>	<input type="checkbox"/>
Other: (Please specify)	<input type="text"/>	<input type="text"/>
Has the patient done their home exercises?	<input type="checkbox"/> Yes <input type="checkbox"/> No (any comments*)	<input type="checkbox"/> Yes <input type="checkbox"/> No (any comments*)

Please record reasons below for any substantial difference from the ProFHER physiotherapy protocol.

Reasons (or comments*)	<input type="text"/>	<input type="text"/>
If referral to other speciality is required, please record speciality and reason	<input type="text"/>	<input type="text"/>

Please continue to record treatment sessions on a new form once this form is complete. A separate form is available to record end of treatment options. Thank you very much for completing this form.

1340011303

Appendix 24 Physiotherapy treatment log: completion of treatment ('end of treatment') form

ProFHER TRIAL: PHYSIOTHERAPY TREATMENT LOG - COMPLETION OF TREATMENT

Participant identification number:

Please record last used sheet number for Physiotherapy Treatment Log:

Name of person completing form:

Please complete appropriate section(s) below.

a) Was treatment completed (formal discharge)? *(Please cross one box)* Yes ☐ No ☐

Date of discharge: / /
Day Month Year

Please record the reason for discharge by crossing one of the following

Independent shoulder function achieved ☐

No improvement noted over several sessions ☐

Other ☐

If Other please describe in box:

b) Was treatment not completed? *(Please cross one box)* Yes ☐ No ☐

Date of discharge: / /
Day Month Year

Please record the reason for discharge by crossing one of the following boxes:

Patient stopped attending ☐

Another problem intervened ☐

If Another problem intervened please describe in box:

Other ☐

If Other please describe in box:

Yes No

c) Did treatment start (i.e. patient never attended)? *(Please cross one box)* ☐ ☐

Please return this and all other physiotherapy treatment log forms whether used or not to the designated person at your hospital. Thank you very much for completing this form.

9204084715

Appendix 25 One-year follow-up form

ProFHER TRIAL: ONE YEAR FOLLOW-UP FORM

Participant identification number:

This form is for the designated person to complete for an eligible patient one year after they were enrolled into the trial.

If the patient has died, please complete and return an adverse event form.

Pre-Discharge:

Has the patient seen an occupational therapist? (Please cross one box)

Yes ☐ No ☐

Date of orthopaedic discharge:
(i.e. end of orthopaedic treatment)

<input type="text"/>	<input type="text"/>	/	<input type="text"/>	<input type="text"/>	/	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Day			Month			Year			

Discharged to: (e.g. GP)
(Please record details)

Complications that occurred after the initial treatment and, if applicable, after inpatient episode:

Please cross in box(es) if recorded

Surgical site infection requiring antibiotics/further surgery

☐

Describe:

Avascular necrosis of humeral head

☐

Describe:

Implant failure

☐

Describe:

Dislocation/instability

☐

Describe:

Metalwork problems requiring further surgery

☐

Describe:

Non-union/Malunion requiring further treatment

☐

Describe:

Other

☐

Describe:

Other treatments after initial treatment and, if applicable, after inpatient episode:

Surgery to shoulder
(please cross one box)

Yes ☐ No ☐

Date:

<input type="text"/>	<input type="text"/>	/	<input type="text"/>	<input type="text"/>	/	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Day			Month			Year			

If 'Yes' describe surgery below.

Number of days as
hospital in-patient

Please continue to complete the rest of this form overleaf.

9753135159

**Increased/new therapy for
shoulder related complication
(e.g. infection)**
(please cross one box)

Yes ☐No ☐Date
started
Day/
Month/
Year

If 'Yes' describe treatment below.

Number of days as
hospital in-patient

**Treatment for serious newly
diagnosed medical
complication**
(please cross one box)

Yes ☐No ☐Date
started:
Day/
Month/
Year

If 'Yes' describe treatment below.

Number of days as
hospital in-patient

Post-Discharge (i.e. end of orthopaedic treatment):

**Has the patient visited the
orthopaedic/fracture clinic
since discharge from
orthopaedic treatment?**
(please cross one box)

Yes ☐No ☐

Date:

Day/
Month/
Year

Number of visits

If 'Yes' please describe the reason(s) for attendance in the box below:

**Has the patient been
admitted to hospital for
another fracture?**
(please cross one box)

Yes ☐No ☐Date
admitted:
Day/
Month/
Year

Number of days as
hospital in-patient

If 'Yes' please provide details in the box below:

Thank you for completing this form.

1035135154

Appendix 26 Two-year follow-up form

ProFHER TRIAL: TWO YEAR FOLLOW-UP FORM

Participant identification number:

This form is for the designated person to complete for an eligible patient between year one and two of being enrolled into the trial.

If the patient has died, please complete and return an adverse event form.

For patients who were not discharged within one year of their original shoulder injury, please record date of orthopaedic discharge:

Date of orthopaedic discharge:
(i.e. end of orthopaedic treatment)

<input type="text"/>	<input type="text"/>	/	<input type="text"/>	<input type="text"/>	/	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Day			Month			Year			

Discharged to: (e.g. GP)
(Please record details)

Complications that happened after one year:

Please cross in box(es) if recorded

Surgical site infection requiring antibiotics/further surgery

☐

Describe:

Avascular necrosis of humeral head

☐

Describe:

Implant failure

☐

Describe:

Dislocation/instability

☐

Describe:

Metalwork problems requiring further surgery

☐

Describe:

Non-union/Malunion requiring further treatment

☐

Describe:

Other

☐

Describe:

Other treatments that happened after one year:

Surgery to shoulder
(please cross one box)

Yes

☐

No

☐

Date:

/

/

Year

If 'Yes' describe surgery below.

Number of days as hospital in-patient

Please continue to complete the rest of the form overleaf.

2320411238

**Increased/new therapy for
shoulder related complication
infection)**

(please cross one box)

Yes ☐

No ☐

**Date
started**

/

/

Day

Month

Year

If 'Yes' describe treatment below.

Number of days as
hospital in-patient

**Treatment for serious newly
diagnosed medical
complication**

(please cross one box)

Yes ☐

No ☐

**Date
started:**

/

/

Day

Month

Year

If 'Yes' describe treatment below.

Number of days as
hospital in-patient

Other patient admissions after one year:

**Has the patient visited the
orthopaedic/fracture clinic this
year?**

(please cross one box)

Yes ☐

No ☐

Date:

/

/

Day

Month

Year

Number of visits

If 'Yes' please describe the reason(s) for attendance in the box below:

**Has the patient been
admitted to hospital for a
fracture this year?**

(please cross one box)

Yes ☐

No ☐

**Date
admitted:**

/

/

Day

Month

Year

Number of days as
hospital in-patient

If 'Yes' please provide details in the box below:

Thank you for completing this form.

8546411234

Appendix 27 Adverse event (reporting) form

ProFHER Trial: Adverse Event Reporting Form

Participant ID number:

This form should be completed for patients who have had an **adverse event**: i.e. any undesirable experience occurring to a patient, whether or not considered related to the treatments being compared in the trial.

Date of assessment (dd/mm/yy):

Name of Principal Investigator:

Name of person reporting the adverse event:

Q1. Please record date of adverse event (dd/mm/yy):

Q2. Nature of adverse event:

Please provide as much information as possible in the box below about the adverse event including investigations and treatment given.

Please continue form overleaf

Q3. Type of adverse event*:

Please categorise this adverse event, crossing all appropriate options:

	Yes	No
a. Resulted in Death	<input type="checkbox"/>	<input type="checkbox"/>
b. Is life-threatening	<input type="checkbox"/>	<input type="checkbox"/>
c. Required hospitalisation or prolongation of existing hospitalisation	<input type="checkbox"/>	<input type="checkbox"/>
d. Resulted in persistent significant disability or incapacity	<input type="checkbox"/>	<input type="checkbox"/>
e. Consists of a congenital anomaly or birth defect	<input type="checkbox"/>	<input type="checkbox"/>
f. Is otherwise considered medically significant by the investigator	<input type="checkbox"/>	<input type="checkbox"/>
g. Other (please describe): <input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q4a. In your opinion, is the adverse event **related** to the patient taking part in the ProFHER Trial?

Yes ☐

No ☐

Q4b. Is it an **expected** ☐ or **unexpected** ☐ adverse event?

Please explain your responses to 4a and 4b in the box below:

* If the answer to any parts of a-f for Question 3 is 'Yes', this is a **serious adverse event** and must be reported to the ProFHER Trial Co-ordinator **within 24 hours** of the local investigator becoming aware of it.

If the only answer to Question 3 is 'Yes' to part g. this is a **non-serious adverse event** and must be reported to the ProFHER Trial Co-ordinator **within 5 days** of the local investigator becoming aware of it.

Please fax this form to the ProFHER Trial Co-ordinator on 01904 [REDACTED] for the attention of Dr Stephen Brealey. Before you fax the form, please ring Stephen Brealey on 01904 [REDACTED] to inform him that the fax is on its way. If Stephen is unavailable then please ring Ms Sarah Gardner, ProFHER Trial Secretary, on 01904 [REDACTED].

Thank you for completing this form

Appendix 28 Review of adverse event form

Proximal Fracture of the Humerus: Evaluation by Randomisation (ProFHER Trial)

REVIEW OF ADVERSE EVENT

Patient concerned
(Trial ID number)

--	--	--	--	--

How, when and by whom
was notification of the
event made

--

Date of review

		/			/				
Day			Month			Year			

Please record what action
was taken to resolve
the adverse event

--

Please record whether
the adverse event has
been resolved

--

Signature of reviewer

--

Date reviewed by Trial
Management Group

		/			/				
Day			Month			Year			

Appendix 29 Trial exit form

ProFHER Trial: Trial Exit Form

Please complete this form when the patient is leaving the trial and should no longer be followed-up

Name of person completing form

Participant identification number:

Exit Date:

<input type="text"/> <input type="text"/>	–	<input type="text"/> <input type="text"/>	–	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
Day		Month		Year

Reason(s) for patient not being followed-up:
(Please place a cross in the appropriate box)

1. Patient wishes to leave the trial

Yes

☐

No

☐

Please state reason (if available).....

2. Patient is being withdrawn by doctor

Yes

☐

No

☐

Please state reason.....

3. Patient is lost to hospital follow-up

Yes

☐

No

☐

Please state reason.....

4. Patient is lost to postal follow-up

Yes

☐

No

☐

Please state reason.....

5. Patient has died

Date of death

<input type="text"/> <input type="text"/>	–	<input type="text"/> <input type="text"/>	–	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
Day		Month		Year

6. Other reason

(Please state below)

Yes

☐

No

☐

.....

Please send this form to the York Trials Unit in the pre-paid envelope provided

Appendix 30 Questionnaire about sling care information

ProFHER TRIAL: QUESTIONNAIRE ABOUT SLING CARE INFORMATION

This section asks about the use of the leaflet that provided advice on initial self-care during sling immobilisation which was intended for all patients who were eligible for the ProFHER trial.

- 1a. Did you give the sling immobilisation leaflet provided in the recruitment pack to patients judged as being **eligible for ProFHER** (whether or not they consented to take part in the trial)?
(Please cross one box only)

All of the time

☐

Most of the time

☐

Rarely

☐

Never

☐

- 1b. If you did not respond 'All of the time' for patients who were **eligible for ProFHER**, please give the reasons for this in the box below.

- 1c. If, for these patients, you routinely used another/other leaflet(s) instead of the sling immobilisation leaflet provided in the recruitment pack could you please describe this/these briefly below and forward a copy of each leaflet to Laura Dennis.

Leaflet type	Please record the reason for having used an alternative leaflet (e.g. a Trust approved leaflet was used)

- 2a. Did you provide a leaflet (either that provided for ProFHER or an alternative) on sling care to patients who **consented to take part** in the trial?
(Please cross one box only)

All of the time

☐

Most of the time

☐

Rarely

☐

Never

☐

- 2b. If you did not respond 'All of the time' for **consenting patients**, please give reasons for this in the box below.

Please continue to complete the rest of the form overleaf.

1360136358

This section asks about the immobilisation that is used for patients who would have been eligible for the ProFHER trial at your hospital.

1. Please describe in the boxes below the type(s) of sling or other methods of immobilisation routinely / typically used for patients who would have been eligible for the ProFHER trial.

Non-surgical treatment:	
Surgically treated:	

2. Typically, how long **in weeks** is the recommended treatment period of arm immobilisation in your hospital for patients who would have been eligible for the ProFHER trial?

Without surgery: weeks

Comment:

Surgically: weeks

Comment:

Please record the name of the person who completed the form:

Please record the name of the hospital:

Thank you very much for your help completing this questionnaire



PROximal Fracture of Humerus: Evaluation by Randomisation (ProFHER) Trial
*A multi-centre randomised controlled trial funded by NHS R&D Health Technology Assessment Programme
 (International Standardised Randomised Controlled Trial Number 50850043)*

5150136359

Appendix 31 Survey of radiography departments

ProFHER TRIAL: Survey of X-ray Departments

- 1a. How many X-ray views are required at your hospital for a patient's first attendance in Accident & Emergency (A&E) with suspected fracture of the proximal humerus?
(Please record the number in the box)

- 1b. Which of the views below are the recommended series of X-rays for these patients?
(Please cross all boxes that apply)

**Anteroposterior
View**

☐

**Axillary
(or modified view)**

☐

**Scapular
Y-Lateral**

☐

If you would like to add any comments, please record these in the box below.

2. The criteria that the two independent orthopaedic surgeons used to assess the quality of X-rays taken for fractures of the proximal humerus are listed below. Can you please record whether you agree or not that these were appropriate criteria.
(Please cross a box for your response to each of the criteria)

Quality criteria:	Agree	Disagree	Record reason you disagree
a. Are there at least two projections in planes perpendicular to each other?	<input type="checkbox"/>	<input type="checkbox"/>	
b. Are the proximal humerus and glenohumeral joint seen on each projection?	<input type="checkbox"/>	<input type="checkbox"/>	
c. Is it possible on the two views to clearly identify the shaft, greater tuberosity, lesser tuberosity, head of humerus and glenohumeral joint?	<input type="checkbox"/>	<input type="checkbox"/>	

0182045415

Please record in the boxes below any other criteria you think should be used for assessing the quality of X-rays for these fractures:

Other quality criteria:	Reason for choosing this:

3. At your hospital, are specialist reports, such as those prepared by a radiologist / radiographer, of the X-rays of the proximal humerus fracture taken in A&E routinely provided to the orthopaedic surgeon for use at the fracture clinic?
(Please cross one box)

All of the time

☐

Most of the time

☐

Rarely

☐

Never

☐

If you would like to add any comments, please record these in the box below.

4. In your opinion, are there any aspects of the interpretation of X-rays of the proximal humerus fractures that are difficult to report on? Please add any comments to this question in the box below:

5480045411

5. In relation to the Neer's classification, which uses both linear displacement and angulation to categorise fracture types, how difficult is it to assess these parameters when viewing X-rays?

(Please cross one box below for each parameter)

	All of the time	Most of the time	Rarely	Never
a. Displacement	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Angulation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

If you would like to add any comments, please record these in the box below.

6. At your hospital, is CT scanning routinely/frequently used for these fractures? Yes ☐

(Please cross one box)

No ☐

If you would like to add any comments, please record these in the box below.

Please record the name of the person who completed the form:

Please record the name of the hospital:

Thank you very much for your help completing this questionnaire.



PROximal Fracture of Humerus: Evaluation by Randomisation (ProFHER) Trial
A multi-centre randomised controlled trial funded by NHS R&D Health Technology Assessment Programme
(International Standardised Randomised Controlled Trial Number 50850043)

6328045413

Appendix 32 Post-recruitment survey of hospital radiographers

Completed surveys on issues relating to imaging and radiograph quality for people with proximal humeral fractures were received from radiographers from 26 participating centres. Eleven (42%) of these were from the trial-designated radiographer listed in the staff delegation log for that centre.

Recommended views

The first two questions of the survey (see *Appendix 31*) asked about the recommended radiographic views for patients attending an A&E department with a suspected fracture of the proximal humerus. Two views were *required* at 25 hospitals and three at one hospital (answers to question 1). *Table 91* shows the responses for the recommended series of radiographs for these patients. Comments on their selection were received from 18 radiographers. The fifth category [anteroposterior view and (axillary view *or* scapular Y-lateral view)] was based on the selection of the views (crossed boxes) and comments from the radiographers on the alternative use of the view (axillary view *or* scapular Y-lateral view) not selected as a recommended view. Of the 13 in this group, nine indicated that the axillary or modified axillary view was the preferred view and one stated that the scapular Y-lateral view was performed more often. Twelve of the 13 in this group and three of the other six radiographers providing feedback commented on the choice of view being dependent on the patient's condition, such as mobility, and related practicalities. Two of these advised that the choice of view would also depend on the presentation of the fracture and two others gave insights on what views were best for dislocation and other fracture types. One referred to additional training being needed for A&E staff and orthopaedic surgeons for interpreting the modified axillary view. One comment referred only to the choice being radiographer dependent.

TABLE 91 Views recommended by hospital radiographers for patients with a proximal humeral fracture ($n = 26$)

Radiographic views	<i>n</i>	%
Anteroposterior and axillary views	6	23
Anteroposterior and scapular Y-lateral views	5	19
Axillary and scapular Y-lateral views	0	0
Anteroposterior, axillary and scapular Y-lateral views	2	8
Anteroposterior view and (axillary view <i>or</i> scapular Y-lateral view)	13	50

Quality criteria

As shown in *Table 92*, the majority of radiographers indicated that they agreed with all three quality criteria used in our interim assessment of baseline radiographs. In all, eight radiographers recorded a disagreement with these criteria, with six disagreeing with one criterion, one disagreeing with two criteria and one disagreeing with all three criteria.

Two of the five comments received on the first criterion pointed out that the modified axial was oblique (45°) to the anteroposterior view; one also noted that the scapular Y-lateral view was oblique to the anteroposterior view. The other three comments related to choice of view, which was covered in the preceding question.

All five comments received on the second criterion referred to potential difficulties in viewing the structures, mainly the glenohumeral joint in the scapular Y-lateral view. These difficulties were restated or implied in five of the six comments on the third criterion. The remaining comment questioned the wording of the criterion, suggesting that it 'should be between the two views that this is shown, not on the two views'.

Eleven radiographers proposed other criteria for assessing the quality of radiographs for proximal humeral fractures. These mainly related to (1) image exposure and contrast and (2) patient positioning.

Provision of specialist reports for use at the fracture clinic

Seventeen radiographers reported that specialist reports, prepared by the radiographer/radiologist, were routinely provided to the surgeon for use at the fracture clinic 'all of the time'. A further seven reported that such provision was 'most of the time' and the remaining two reported that it was 'rarely'. Pertinent comments on this issue related to the timing of the fracture clinic and availability of personnel. One radiographer considered that 'The question is a nonsense with PACS'. (These responses and responses to later questions implied that, although the specialist report may have been available it was unlikely that it informed the interpretation of the radiographs by the surgeon in terms of patient eligibility for the ProFHER trial.)

Perceived difficulties in interpretation of radiographs of the proximal humerus

The survey asked about difficulties in interpretation in a general way and then specifically in relation to Neer's linear displacement and angulation criteria. Of the 20 responses to the general question, eight implied that there were difficulties resulting from poor-quality images, relating, according to six radiographers, to one or all of the following: patient positioning/condition, inadequate exposure and poor technique. Four responses referred to limitations of or challenges of interpreting specific radiographic views, specifically scapular Y-lateral and the modified axillary views. One response observed that A&E

TABLE 92 Radiographer agreement with quality criteria ($n = 26$)

Radiographic views	Agree with criterion, n	%
1. Are there at least two projections in planes perpendicular to each other?	24	92
2. Are the proximal humerus and glenohumeral joint seen on each projection?	22	85
3. Is it possible on the two views to clearly identify the shaft, greater tuberosity, lesser tuberosity, head of the humerus and glenohumeral joint?	21	81

referral often resulted in the selection of the less useful anteroposterior image (presumably in the coronal rather than the scapular plane) for interpreting proximal humeral fractures. Six radiographers referred to limitations or challenges in relation to the fracture itself. These related to more complex and severe fractures in four responses and to 'subtle' or undisplaced fractures in the other two responses. As captured in the final pithy response ('Generally not'), there was no indication that there were problems in interpretation that made reporting of these fractures routinely difficult for radiographers.

Of the 17 responses on difficulties in interpretation relating to Neer's linear displacement and angulation criteria, nine indicated that it was difficult most of the time and eight indicated that it was rarely difficult. Of the 12 comments on this aspect, six emphasised that they did not use the Neer classification and one other commented on the known problems regarding inter-rater agreement for this classification. Two radiographers remarked that the question was difficult to understand; one of these though suggested that it was usually possible to assess both linear displacement and angulation. One radiographer's response gave this impression too but in the context of image quality, whereas another radiographer suggested that reporting of these was 'sometimes difficult'. The final response referred to the need for a scapular Y-lateral view to check for displacement.

On review of these responses, which gave no real insights into the quantification of fracture displacement, we realised that our emphasis on the Neer classification had been unhelpful.

Routine use of computerised tomography scans

One radiographer reported that CT scanning was 'routinely/frequently used for these fractures' in her centre and 24 reported that it was not (the answer was missing for one centre). Comments received from 10 radiographers reinforced this picture, with some reference made by four radiographers to the use of CT to augment diagnosis subsequently in the case of poor-quality radiographs and/or complex fractures.

Appendix 33 The Neer classification of radiographs pro forma

ProFHER TRIAL: Neer's X-ray Classification Proforma

Trial Participant ID : Rater :

1. Please record a cross in the box(es) below that apply:

X-ray ID	Anteroposterior (Scapular plane)	Anteroposterior (Coronal plane)	Axillary (or modified axillary)	Scapular Y-lateral	Other - please describe
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

2. Are there at least two projections in planes perpendicular to each other?
(Please cross one box only) Yes ☐ No ☐

3. For each X-ray projection is the:

X-ray ID	Proximal humerus visible?	Glenohumeral joint visible?	Clip on Left / Right marker available? *	Clear / gross distortion in size?
	Yes / No	Yes / No	Yes / No	Yes / No
<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>

* This applies to a clip that was added to the cassette at the time of image acquisition

4. Are proximal humerus and glenohumeral joint seen on each projection
[of at least two different views?] (Please cross one box only) Yes ☐ No ☐

5. For each X-ray, can you clearly identify the following:

X-ray ID	Humeral shaft?	Greater tuberosity?	Lesser tuberosity?	Head of humerus?	Glenohumeral joint?
	Yes / No	Yes / No	Yes / No	Yes / No	Yes / No
<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>

6. Considering all the available views together, can you visualise the location of all five structures (the humeral shaft, greater tuberosity, lesser tuberosity, head of humerus and glenohumeral joint) sufficiently to determine the position and displacement of the fractured segments? (Please cross one box only) Yes ☐ No ☐

7. Please give your overall assessment of quality for this set of images (i.e. in terms of exposure and patient positioning) based on the three grades: (Please cross one box only)

- a) Good: Optimal image definition & content, good for classification
b) Fair: Suboptimal image definition & content, but adequate for classification
c) Poor: Suboptimal image definition & content, inadequate for classification

☐
☐
☐

4291527808

Displacement

Criteria / terms for displacement:

- "involved" - this includes any indication of a fracture for that anatomical structure
- "undisplaced" - either not displaced or clearly insufficient to meet Neer's criteria
- "displaced (unclear if Neer)" - indicates marginal cases: some uncertainty that displacement meets Neer's criteria (lack of scale marker or poor quality image may be reasons)
- "displaced (Neer)" - gross (including complete displacement), clearly meets Neer's criteria (45°; 1cm - linear measurement may not be possible for some x-rays)

"No contact fracture" (of the surgical neck) - the shaft has displaced fully in relation to the head and these segments have no contact with each other

Fracture dislocation applies to the dislocation of the humeral head from the glenoid.

8. Please record a cross in the box(es) below that apply: (N.B. If you do record that an anatomical structure is 'involved' then choose one of the three other descriptors (undisplaced, displaced (unclear if Neer), displaced (Neer)) to indicate the extent of displacement according to Neer's criteria.

	X-ray ID <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	X-ray ID <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	X-ray ID <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	X-ray ID <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	Verdict
Anatomical neck					
- involved	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- undisplaced	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- displaced (<u>unclear if Neer</u>)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- displaced (<u>Neer</u>)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Surgical neck					
- involved	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- undisplaced	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- displaced (<u>unclear if Neer</u>)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- displaced (<u>Neer</u>)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- no contact fracture	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- impacted	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Greater tuberosity					
- involved	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- undisplaced	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- displaced (<u>unclear if Neer</u>)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- displaced (<u>Neer</u>)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Lesser tuberosity					
- involved	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- undisplaced	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- displaced (<u>unclear if Neer</u>)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- displaced (<u>Neer</u>)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fracture dislocation					
- anterior	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- posterior	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Head splitting #					
Is there an articular surface fracture?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Head position					
Head segment in varus	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Head segment in valgus	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9. Using the separate sheet provided of Neer's classification 1 to 16 and your verdicts above ("displaced (Neer)" only), please record Neer's category:

☐☐

9049527804

Appendix 34 List of hospitals contributing to the ProFHER trial and primary care trusts for which research and development approval was obtained for follow-up

Site ^a	Hospital name and/or trust (if applicable)
Basildon	Basildon Hospital, Basildon and Thurrock University Hospitals NHS Foundation Trust
Basingstoke	Basingstoke and North Hampshire Hospital
Birmingham	Queen Elizabeth Hospital Birmingham, University Hospitals Birmingham NHS Foundation Trust
Blackburn	Royal Blackburn Hospital, East Lancashire Hospitals NHS Trust
Blackpool	Blackpool Victoria Hospital, Blackpool, Fylde and Wyre Hospitals NHS Foundation Trust
Coventry	University Hospitals Coventry and Warwickshire NHS Trust
Fife	Queen Margaret Hospital (Dunfermline), NHS Fife
Frenchay	Frenchay Hospital, North Bristol NHS Trust
Glan Clwyd	Ysbyty Glan Clwyd Hospital (Rhyl), North Wales NHS Trust
Ipswich	Ipswich Hospital NHS Trust
James Cook	James Cook University Hospital (Middlesbrough), South Tees Hospitals NHS Foundation Trust
Kings	King's College Hospital NHS Foundation Trust (London)
Leeds	Leeds General Infirmary, Leeds Teaching Hospitals NHS Trust
Leicester	Leicester General Hospital, University Hospitals of Leicester NHS Trust
Lincoln	Lincoln County Hospital, United Lincolnshire Hospitals NHS Trust
Liverpool	Royal Liverpool University Hospital, Royal Liverpool and Broadgreen University Hospitals NHS Trust
Manchester	Manchester Royal Infirmary, Central Manchester University Hospitals NHS Foundation Trust
Newcastle upon Tyne	Freeman Hospital, Newcastle upon Tyne Hospitals NHS Foundation Trust
Norfolk	Norfolk and Norwich University Hospitals NHS Foundation Trust
North Staffordshire	University Hospital of North Staffordshire NHS Trust (Stoke-on-Trent)
North Tees	University Hospital of North Tees (Stockton-on-Tees), North Tees and Hartlepool Hospitals NHS Foundation Trust
North Tyneside	North Tyneside General Hospital (North Shields)
Northampton	Northampton General Hospital NHS Trust
Nottingham	Queen's Medical Centre, Nottingham University Hospitals NHS Trust
Oldham	Royal Oldham Hospital, Pennine Acute Hospitals NHS Trust
Oxford	John Radcliffe Hospital, Oxford Radcliffe Hospitals NHS Trust
Royal London	Royal London Hospital (London), Barts and the London NHS Trust
Salford	Salford Royal Hospital, Salford Royal NHS Foundation Trust
Scunthorpe	Scunthorpe General Hospital, North Lincolnshire and Goole Hospitals NHS Foundation Trust

Site ^a	Hospital name and/or trust (if applicable)
Southport	Southport and Ormskirk District General Hospital, Southport and Ormskirk Hospital NHS Trust
Sunderland	Sunderland Royal Hospital, City Hospitals Sunderland NHS Foundation Trust
Torquay	Torbay Hospital, South Devon Healthcare NHS Foundation Trust
Trafford	Trafford General Hospital (Manchester), Trafford Healthcare NHS Trust
Wansbeck	Wansbeck General Hospital (Ashington), Northumbria Healthcare NHS Foundation Trust
Yeovil	Yeovil District Hospital, Yeovil District Hospital NHS Foundation Trust

^a These were the names used in the trial for the individual sites. With some notable exceptions, most correspond to the location of the hospital. When there are exceptions, the hospital location has been stated in the next column.

Primary care trusts

Site	Primary care trust
Basingstoke	Hampshire Community Health Care
Blackburn	NHS East Lancashire NHS Blackburn with Darwen
Leicester	NHS Leicestershire County and Rutland NHS Leicester City
Newcastle upon Tyne	NHS Newcastle and North Tyneside Community Health
Norfolk	NHS Norfolk NHS Great Yarmouth and Waveney
Nottingham	NHS Nottinghamshire County
Oxford	Thames Valley Primary Care Research Partnership
Royal London	NHS Tower Hamlets
Salford	NHS Salford
Torbay	NHS Devon NHS Torbay Care Trust
Yeovil	NHS Dorset Primary Care Trust Taunton & Somerset R&D Consortium

Appendix 35 Health-care collaborators at participating acute trust centres

Health-care collaborators have been identified from delegation logs and site-specific information documents.

Centre	Principal investigator	Named surgeon – involved in assessing eligibility/form completion	Study co-ordinator (e.g. research nurse/physiotherapist)	Physiotherapy contact	Radiography contact
Basilston and Thurrock University Hospitals NHS Foundation Trust, Basilston Hospital, Nethermayne	Mr Tim Peckham	Mr Ihsan Ahmed	Mrs Nhlanhla Mguni, Mrs Lesley Cruse, Mr Maxwell Masuka	Mrs Usharani Matthias	
	Mr Nigel Rossiter	Mr John Britton, Mr Geoff Stranks, Mr Jonathan Hobby, Mr Kevin Conn, Mr James Calder, Mr Richard Harker, Mr Adrian Wilson, Mr Dominic Simon, Mr Vasselin Vratichowski	Mrs Mary Wright	Mr Kevin Spear, Mrs Jennie Leech	Mr John Beamer
University Hospitals Birmingham NHS Foundation Trust Queen Elizabeth Hospital Birmingham	Mr Socrates Kalogrianitis	Mr Vasilis Tsouparopoulos	Mrs Sarah Holloway, Mrs Nicola Anderson	Mrs Catherine Burford (née Cooke)	Mrs Helen Fowler
	Mr Bambos Charalambous	Mr Vijay Kamath, Mr Stephen McLoughlin, Mr Stephen Mannion	Mrs Sally Baddeley, Mr Stephen Preston, Mrs Lesley Rice	Mrs Clare Blundell	Mrs Susan Field
	Mr Mark Crowther	Mr Iain Packham, Mr Mike Kelly	Mrs Rebecca Fox, Mr Steven Bamfield, Mrs Ruth Halliday	Ms Rhian Witham	Mr Ross Cook
University Hospitals Coventry and Warwickshire NHS Trust Clinical Sciences Research Unit, Warwick Medical School University of Warwick, Warwick	Mr Matthew Costa	Mr Steve Drew, Mr Martin Blakemore, Mr Stephen Turner, Mr David Wainwright	Mrs Helen Richmond, Mr Phillip Jones, Mrs Rebecca Kearney, Mr Troy Douglin, Mrs Kate Dennison, Mrs Catherine Richmond, Mrs Katie Dunn (née McGuinness), Dr Juul Atchen		
NHS Fife, Queen Margaret Hospital Orthopaedics Department, Dunfermline	Mr Sunil Sharma	Mr Robert Cook, Mr Ivan Brenkel	Mrs Lynne Gordon, Mrs Fiona Cameron		

Centre	Principal investigator	Named surgeon – involved in assessing eligibility/form completion	Study co-ordinator (e.g. research nurse/physiotherapist)	Physiotherapy contact	Radiography contact
Ipswich Hospital NHS Trust Department of Orthopaedics, Ipswich	Mr Christopher Roberts	Mr Paul Crossman	Mrs Claire Tricker, Mrs Jenny Finch	Mrs Stephanie Walker, Mr Mark Alderton, Mr Theo Latham	
King's College Hospital NHS Foundation Trust, London	Mr Joydeep Sinha	Mr Anand Arya, Mr Adel Tavakkolizadeh	Mrs Sarah Brannigan	Mr Stephen Daly	Mr Martin Shute
Leeds Teaching Hospitals NHS Trust, Leeds General Infirmary Orthopaedics Department, Leeds	Mr Roger Hackney	Mr George Kontakis, Mr Adam Rumian, Mr Simon Fogerty, Mr Jeremy Viner, Mr Neil Patel	Mrs Jennifer Ogden	Mrs Fiona Watts, Mrs Lorna Williams	
University Hospitals of Leicester NHS Trust, Leicester General Hospital, Leicester	Miss Alison Armstrong	Mr Amit Modi, Mr Radhakant Pandey, Mr Grahame Taylor	Mrs Emma Alderman	Mrs Helen Fort	Mrs Joanne Wormleighton
United Lincolnshire Hospitals NHS Trust, Lincoln County Hospital, Lincoln	Professor Mohammad Maqsood	Mr Latif Mubasher, Mr David Gale, Mr Mark Roswell, Mr Michael Feeney, Mr Ramesh Najak, Miss Saj Hussain, Mr Senbaga Needhirajan, Mr Rajan Asinivatham, Mr Saeed Qaimkhani, Mr Christopher Lee, Mr Abdul Baco	Mrs Karen Warner, Ms Alison Raynor	Mr Mark Fielding, Mr John Wilson, Mr Gopi Reddy	Mr Mujahid Kamal
Central Manchester University Hospitals NHS Foundation Trust, Manchester Royal Infirmary, Manchester	Mr Dilraj Sandher	Mr Nicholas Kenny	Mrs Fiona Holmes, Mrs Andrea Byrne	Mrs Justine Theaker	Mr Jonathan Kitching
Newcastle upon Tyne Hospitals NHS Foundation Trust, Freeman Hospital, Newcastle upon Tyne	Mr Andrew Gray	Mr John Williams, Mr Peter Worlock, Mr Paul Fearon	Mrs Jacqueline Claydon, Mrs Alison Steele	Mrs Nicola Rose, Mrs Karthikeyan Muthumayandi	Mrs Louise Fox

Centre	Principal investigator	Named surgeon – involved in assessing eligibility/form completion	Study co-ordinator (e.g. research nurse/physiotherapist)	Physiotherapy contact	Radiography contact
Norfolk and Norwich University Hospitals NHS Foundation Trust Department of Orthopaedic Surgery, Norwich	Professor Simon Donell	Mr Ben Davis, Mr Adrian Chojnowski, Mr Peter Hallam, Mr Amratlal Patel	Mrs Sarah Price, Mrs Tracey Potter, Mrs Adele Cooper, Mrs Elizabeth Saunders, Mrs Sarah Ware, Mrs Loretta Dean, Mrs Melissa Cambell-Kelly	Mr Toby Smith	Mrs Karen Reid, Mrs Sarah Coward
University Hospital of North Staffordshire NHS Trust Department of Trauma and Orthopaedics, Stoke-on-Trent	Mr Damian McClelland		Mrs Natalie Grocott, Mrs Sarah Griffiths		Mr Praveen Datta
North Tees and Hartlepool Hospitals NHS Foundation Trust, Stockton-on-Tees	Mr Alan Middleton	Mr Neil Bayliss	Mrs Lesley Boulton	Mr Atle Karstad	Mr Matthew Trehella
North Tyneside General Hospital, North Shields	Mr Roland Pratt	Mr Seif Asaad, Mr David Cloke, Mr Dominic Inman, Mr Malcolm Scott, Mr Simon Jones	Mrs Angela Larkin		Mrs Deborah Henderson
Northampton General Hospital NHS Trust, Northampton	Mr Alistair Jepson	Mr Gregor Kerr, Mr David Gidden	Mrs Susan Brown	Mrs Dakshika Govan, Mr Vijay Rajaram	Mr John Middleton
Nottingham University Hospitals NHS Trust Division of Orthopaedic and Accident Surgery, Queen's Medical Centre Campus, Nottingham	Professor Angus Wallace	Mr Paul Manning, Mr John Geoghegan, Mr Giles Becker, Mr Tom Rowlands, Mr Greg Heffernan	Mrs Alison Raynor, Mr Pete Johnson, Mr Andrew Dainty, Dr Litza Vera	Mrs Catherine Clarke, Mr David Whitaker	
Pennine Acute Hospitals NHS Trust, Royal Oldham Hospital, Oldham	Mr Leo Jacobs, Mr Robin Seagger		Mrs Lorraine Lock	Mrs Diane Lewis, Mrs Catherine Maloney	Mr Tim Barratt
Oxford Radcliffe Hospitals NHS Trust, Kadoorie Centre for Critical Care Research John Radcliffe Hospital, Oxford	Professor Keith Willett	Mrs Rachel Denyer, Mr Joseph Alsousou	Mrs Bridget Gray, Mrs Susanna Symonds, Mrs Elizabeth Oaster		

Centre	Principal investigator	Named surgeon – involved in assessing eligibility/form completion	Study co-ordinator (e.g. research nurse/physiotherapist)	Physiotherapy contact	Radiography contact
East Lancashire Hospitals NHS Trust, Royal Blackburn Hospital, Blackburn	Mr Makaram Srinivasan	Mr Hans Marynissan, Mr Kevin Sharpe	Mrs Helen Thompson, Mrs Gillian Haworth, Mrs Rachel Dean, Mrs Julie Grundy, Mrs Adrienne Whittle		Mr David Enion
Royal Liverpool and Broadgreen University Hospitals NHS Trust, Department of Orthopaedics, Broadgreen Hospital, Liverpool	Mr Matthew Smith	Mr Inigo Guisasola	Mrs Fiona Cowell, Mrs Katy Clay, Mrs Paula Houghton		Mr Steven McDonald
Barts and the London NHS Trust, Royal London Hospital, London	Mr Simon Owen-Johnstone	Miss Ang SweetChai	Mrs Gayle Maffulli	Ms Susan Sandford-Smith	Ms Rachael Bowie
Salford Royal NHS Foundation Trust, Department of Trauma and Orthopaedics, Salford Royal Hospital, Salford	Mr Muthu Jeyam		Mrs Jane McConniffe, Mrs Sophie Campbell	Mrs Kathleen Tatlow	
North Lincolnshire and Goole Hospitals NHS Foundation Trust, Department of Trauma and Orthopaedics, Scunthorpe General Hospital, Scunthorpe	Mr Owais Shafqat		Mrs Dianne Briggs	Mrs Susan Batley, Mrs Linda Weaver	
Southport and Ormskirk Hospital NHS Trust, Southport and Ormskirk District General Hospital, Southport	Mr William Loh, Mr Jayant Nadkarni	Mr Kartik Iyengar	Mrs Beryl Thompson, Mrs Anna Morris	Mrs Beverly Webster, Mr Jonathan Miller, Mrs Kimberley Manktelow	Mr David Lodwig, Mrs Melanie Bodenhan
City Hospitals Sunderland NHS Foundation Trust, Sunderland Royal Hospital, Sunderland	Mrs Lisa Tourett		Mrs Julie Walmsley	Mrs Paula Jordan, Mr David Magee	Mr John Parker
South Tees Hospitals NHS Foundation Trust, James Cook University Hospital, Middlesbrough	Mr Amar Rangan	Mrs Maya Rookmoneea, Mr Raymond Liow	Mrs Lucy Micklewright, Mrs Birgit Hanusch, Mrs Lucksy Kottam, Mrs Rachel Clarkson	Mrs Lorna Goodchild	

Centre	Principal investigator	Named surgeon – involved in assessing eligibility/form completion	Study co-ordinator (e.g. research nurse/physiotherapist)	Physiotherapy contact	Radiography contact
South Devon Healthcare NHS Foundation Trust, Torbay Hospital, Torquay	Miss Veronica Conboy	Mr Rangaraja Ramesh	Mr Andrew Hall, Mrs Pauline Mercer, Mrs Gabrielle de Selincourt	Mrs Rebecca Ainsworth	Mrs Michelle Reed, Mrs Emma Jones, Mr Timothy Simpson
Trafford Healthcare NHS Trust, Department of Trauma and Orthopaedics, Trafford General Hospital, Manchester	Mr Bibhas Roy		Mrs Sophie Campbell	Mrs Deborah Loughhead	
Northumbria Healthcare NHS Foundation Trust, Wansbeck General Hospital, Ashington	Mr Jaime Candal-Couto	Mr Chris Gibbons, Mr Simon Jones, Mr Scott Muller, Mr Aradhyula Murty	Mrs Caroline Clark		
Yeovil District Hospital NHS Foundation Trust, Department of Trauma and Orthopaedics, Yeovil District Hospital, Yeovil	Mr Andrew Chambler	Mr John Smibert, Mr Amitab Dwyer, Mr Tarek El-Bouni	Mrs Barbara Williams-Yesson, Miss Clare Buckley, Mrs Rebecca Rowland-Axe	Mr Matthew Davey	
North Wales NHS Trust, Ysbyty Glan Clwyd Hospital, Rhyl	Mr Amit Sinha	Mr Raghavendra Madhusudan, Mr Sanath Kumar Shetty, Mr Ahmed Humayun Mirza, Mr Balasundaram Ramesh	Mr Raghavendra Madhusudan, Mr Sanath Kumar Shetty	Mr Krishnan Prabhakaran	Mrs Bethan Wyn Owen

Appendix 36 Report of the processing of complications and associated treatment including 'discards'

With interim feedback and checks from the statistician (AK) to ensure consistency, two raters screened and coded data from a file with 273 entries for complications and treatments drawn from the relevant fields for each trial patient from three forms (end of orthopaedic inpatient episode form, 1-year follow-up form and 2-year follow-up form) (see *Chapter 3*). One rater (HH) coded the 'discards' into eight categories, which were honed down to four categories for reporting purposes by AK: repeat information, subsequent/additional information to treatment for a complication, irrelevant content and, for further fracture data, routine follow-ups. Additionally, incorrectly reported categories were reclassified and information from another source (these were surgery forms received for two non-surgical patients who received surgery because of an early complication – these data were handled as if they had been reported at the 1-year follow-up; coding was by the two raters) was identified and added to the definitive data set. Entries were also added for surgical and shoulder fracture-related complications that were implied only from descriptions of treatment.

The following gives an account of the 'discards' and adjustments for the various categories of complications.

Surgical and shoulder fracture-related complications

A total of 81 surgical and shoulder fracture-related complications were reported over the 2-year follow-up period. In total, 34 of these reports were discarded (23 contained repeat information, six were subsequent/additional information to treatment for a complication and five contained irrelevant content). Of the remaining 47 complications, two had been assigned a wrong category on the hospital form and were recategorised. A total of 14 complications were added (one recategorisation from a different section on a form and 13 additional complications derived from other available information). The final data therefore consisted of 59 surgical and shoulder fracture-related complications reported for 53 patients.

Of the 48 treatment reports (including the two reported in surgery forms only) for these complication, 13 were discarded (five contained repeat information, seven were subsequent/additional information to treatment for a complication and one contained irrelevant content). Of the remaining 35 treatments, two had been assigned a wrong category and were recategorised, leaving 33 valid entries reported for 33 patients.

Medical complications during the inpatient stay

A total of eight medical complications were reported, to which a further two were added that had been entered under the wrong category on other forms, resulting in a total of 10 valid medical complications for 10 patients.

Treatment for serious newly diagnosed medical complication

A total of 64 treatments for medical complications were reported over the 2 year follow-up period. In total, 12 of these reports were discarded (11 contained repeat information and one contained irrelevant content). Of the remaining 52 treatments, one had been assigned a wrong category and was recategorised, leaving 51 valid treatments for 46 patients.

Further fractures

Of a total of 23 reported hospital admissions for another fracture, five were discarded (repeat information), resulting in 18 valid entries for 18 patients.

Of a total of 51 reported visits to orthopaedic/fracture clinics, 27 were discarded (seven contained repeat information, nine were subsequent/additional information to treatment for a complication, eight related to routine follow-ups and three contained irrelevant content). To the remaining 24 visits, two were added that had been assigned the wrong category in other sections of the form, resulting in a total of 26 valid visits for 25 patients.

Appendix 37 Additional sensitivity analysis exploring the effect of clustering at centre level


Feedback obtained subsequent to the preparation of this report prompted a formal analysis to examine the effect of clustering at centre level. This is in addition to our descriptive analysis of centre variability presented in *Chapter 5*.

We conducted an extra sensitivity analysis whereby centres were added as a random effect to the primary model to explore the effect of clustering at the centre level. Adjusted OSS means and group differences for this model are presented in *Table 93*. Similar to the primary analysis, there was no significant overall effect of treatment group over the 2-year follow-up period and no significant effect of treatment group at individual time points. Therefore, the results of the primary analysis are robust to the effects of clustering by centre. As with the primary and all subgroup and sensitivity analyses, all 95% CIs for the treatment effect excluded the prespecified difference of 5 OSS points, representing 'clinical significance'.

TABLE 93 Difference in mean OSS scores over time by treatment group: sensitivity analysis – adjusted for clustering by centre^a

Follow-up	Surgery (<i>n</i> = 125) (<i>n</i> = 114 in analysis), mean (95% CI)	Not surgery (<i>n</i> = 125) (<i>n</i> = 117 in analysis), mean (95% CI)	Difference (95% CI)	<i>p</i> -value
Overall	39.06 (37.18 to 40.86)	38.27 (35.81 to 40.03)	0.79 (−1.30 to 2.88)	0.460
6 months	37.83 (35.81 to 39.75)	35.54 (33.48 to 37.49)	2.29 (−0.03 to 4.61)	0.053
12 months	39.22 (37.26 to 41.08)	38.76 (36.86 to 40.57)	0.46 (−1.75 to 2.66)	0.686
24 months	40.10 (38.13 to 41.99)	40.36 (38.48 to 42.16)	−0.26 (−2.51 to 1.99)	0.821

^a Multilevel model accounting for centre as a random effect, adjusted for treatment group, time (6, 12 and 24 months), group × time interaction, baseline EQ-5D index, gender, age group (< 65 years/≥ 65 years) and tuberosity involvement at baseline (yes/no).

A decorative graphic consisting of numerous thin, parallel green lines that curve from the left side of the page towards the right, creating a sense of movement and depth.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health

Published by the NIHR Journals Library